Assortative interactions revealed by sorting of animal groups

Alex Szorkovszky[1*], Alexander Kotrschal[2*], James E Herbert Read[2], Severine D Buechel[2], Maksym Romenskyy[1], Emil Rosén[1], Wouter van der Bijl[2], Kristiaan Pelckmans[3], Niclas Kolm[2], and David JT Sumpter[1]

1. Mathematics Department, Uppsala University, Sweden

2. Zoology Department, Stockholm University, Sweden

3. IT Department, Uppsala University, Sweden

* These two authors contributed equally to this work.

Corresponding author:

Alex Szorkovszky

alexander.szorkovszky@math.uu.se

Ph: +467 3892 2025

Mathematics Department

Box 480

75106 Uppsala

Sweden

Word count: 5311 + references and captions

Funding

This work was supported by the Knut and Alice Wallenberg Foundations, grant number 102 2013.0072.

Abstract

Animals living in groups can show substantial variation in social traits and this affects their social organisation. However, as the specific mechanisms driving this organisation are difficult to identify in already-organised groups typically found in the wild, the contribution of inter-individual variation to group-level behaviour remains enigmatic. Here, we present results of an experiment to create and compare groups that vary in social organisation, and study how individual behaviour varies between these groups. We iteratively sorted individuals between groups of guppies (*Poecilia reticulata*) by ranking the groups

according to their directional alignment and then mixing similar groups. Over the rounds of sorting the consistency of the group rankings increased, producing groups that varied significantly in key social behaviours such as collective activity and group cohesion. The repeatability of the underlying individual behaviour was then estimated by comparing the experimental data to simulations. At the level of basic locomotion, individuals in more coordinated groups displayed stronger interactions with the centre of the group, and weaker interactions with their nearest neighbours. We propose that this provides the basis for a passive phenotypic assortment mechanism that may explain the structures of social networks in the wild.

Keywords: collective behaviour, repeatability, sociability

Group living can reduce predation risk (Hamilton, 1971; Seghers, 1974; Foster & Treherne, 1981; Magurran & Seghers, 1994), improve reproductive opportunities (Krause & Ruxton, 2002; Silk, 2007) and provide access to social information about the location of food and shelter (Sumpter & Pratt, 2009; Sumpter, 2010; Pike & Laland, 2010; Miller, Garnier, Hartnett, & Couzin, 2013). However, groups of animals are typically not behaviourally uniform. Individuals of the same species commonly differ in repeatable inter-individual behaviour, also known as behavioural phenotypes, for traits such as boldness, aggression and sociability (Réale, Reader, Sol, McDougall, & Dingemanse, 2007; Sih, Cote, Evans, Fogarty, & Pruitt, 2012; Wolf & Weissing, 2012).

How these traits affect social organisation and therefore impact group behaviour is still not completely understood. Behavioural phenotypes can affect the function and organisation of groups in at least three ways. First, properties of the group that emerge from many inter-individual interactions can be affected by the presence or absence of different behavioural types in the group, that is, on its "group phenotypic composition" (Farine, Montiglio, & Spiegel, 2015). For instance, more variation in boldness affects the shape of animal groups (Couzin, Krause, James, Ruxton, & Franks, 2002; Killen, Marras, Nadler, & Domenici, 2017) and their spatial distribution (Michelena, Jeanson, Deneubourg & Sibbald 2010). On longer time scales, the composition of behavioural types affects the survival of groups, and hence this may be subject to selection (Pruitt & Goodnight,

2014). Second, behaviour of the individuals within the group may also depend on the behavioural phenotypic composition of the group (Webster & Ward, 2011; Dingemanse & Araya-Ajoy, 2015). For example, conformity to the average group behaviour is widely observed (Herbert-Read et al, 2013; King, Williams, & Mettke-Hofmann, 2015), and the resulting similarity across group members can reduce risk of predation (Landeau & Terborgh, 1986). Certain behaviours may also be expressed to compensate for a lack of variation in a group, for instance by modulating aggression in order to reduce risk in conflict (Sih & Watters, 2005). Third, it has been shown that individuals can actively associate with other individuals depending on their phenotypes (Krause, Butlin, Peuhkuri, & Pritchard, 2000). For example, associating with dissimilar behavioural phenotypes may confer an advantage for competitive foragers (Metcalfe & Thomson, 1995).

These three mechanisms (which we will refer to respectively as emergence, behavioural plasticity and active self-assortment) are functionally distinct, but can all lead to animal groups being structured according to behaviour, which poses a challenge for inferring which mechanism applies. In addition, it is often difficult to analyse consistent differences between groups in the wild, such as when group membership is constantly changing. One fruitful mode of observational study has been social network analysis, in which the strength of social ties between pairs of individuals may be quantified by propensity to co-occur in the same groups (Croft et al., 2005; Sundaresan, Fischhoff, Dushoff, & Rubenstein, 2007; Farine & Whitehead 2015). These can be used to infer that, for example, individuals self-assort by shoaling tendency, as reflected in the structure of the network (Croft et al., 2005). However, using such methods, the role of inter-individual influence on individual social behaviour still cannot be ruled out (Shalizi & Thomas, 2011). To determine the role of social context, laboratory methods may be used, such as analysing responses to specific phenotypic compositions (Magnhagen & Staffan, 2005; Pike, Samanta, Lindström, & Royle, 2008; Dyer, Croft, Morrell, & Krause, 2009).

Here we use a novel method in a laboratory setting to maximise the variation in shoaling tendency between groups of guppies (*Poecilia reticulata*) that resembles the variation between self-assorted

groups in the wild. By creating groups with consistent differences in individual behaviour, we can investigate the traits underlying the properties of shoals, and hence obtain insights into how specific group behaviours may evolve under selection (Ioannou, Guttal, & Couzin 2012). Guppies are a model species in the study of anti-predator shoaling behaviour (Farr, 1975; Dugatkin & Godin, 1992; van der Bijl, Thyselius, Kotrschal, & Kolm, 2015; Herbert-Read et al., 2017), known for fission-fusion dynamics and self-assortment according to sociability (Croft et al., 2005). We investigate the differences between these sorted groups' shoaling behaviours, to identify possible mechanisms for self-assortment. We divide three independent collections of 128 guppies each into 16 groups of 8. We subject each of these groups repeatedly to open field assays to quantify their directional 'alignment' — that is, the degree to which the eight guppies are moving in the same direction. This measurement combines cohesiveness, crucial in the 'selfish herd' response to predation (Hamilton, 1971), and co-ordination, which facilitates information transfer (Strandburg-Peshkin et al., 2013; Rosenthal, Twomey, Hartnett, Wu, & Couzin, 2015). In subsequent iterations, we manipulate the membership of the groups based on the results of the previous round (i.e. we switch individuals between groups that show similar social scores).

Assuming that variation in group alignment is primarily driven by behavioural phenotype, we predict that groups keep consistent rankings between rounds. We further predict that with an increasing number of iterations, this consistency will increase as the traits become sorted according to phenotype, and hence, within-individual variation becomes relatively less important. By fitting our experimental results to a simulated model of the sorting process, we estimate the trait repeatability R. Finding R from purely group-level data may seem counter-intuitive, but maximum-likelihood fitting is possible due to the fact that the sorting dynamics depend heavily on the underlying variation between individuals (Szorkovszky et al., 2017). We then analyse differences between the sorted groups at three scales: at the group level, at the level of subgroups (local aggregations) and at the level of basic locomotion and interactions. Using this data, we then demonstrate how variation at all levels may provide a mechanism for self-assortment as seen in the

wild.

Methods

Sorting

The laboratory population of guppies used for this study originated from a down-stream population of the Quare river in Trinidad, subject to high predation levels. The original collection was made in 1998 (Pélabon et al., 2014) and the lab population has since been kept in several large (>500 litre) tanks of >500 individuals each to avoid inbreeding. All described experiments here were performed in a sub-set of this original collection at the Stockholm University aquatic facilities. The laboratory was maintained at 26°C with a 12:12 light:dark schedule. Fish were fed a diet of flake food and freshly hatched brine shrimp six days per week.

[FIGURE 1 AROUND HERE]

On the first day of filming, mature, unmarked female guppies were allocated into 16 groups of eight fish such that all conspecifics within each group were unfamiliar to each other. Each group was kept in a seven litre tank containing two cm of gravel and a biological filter. We allowed for visual contact between the tanks. After this initial allocation, we used an iterative procedure of mixing groups to sort the guppies. In each round of the experiment, 16 new groups were created and immediately filmed in a shoaling assay. These videos were then tracked and each group was evaluated before the following round. The first round was filmed without sorting and, from the second round onwards, the groups were paired and mixed according to their rankings in the previous round. For instance, if a group was ranked 8th in the previous round we would exchange four fish from this group with four fish from the group ranked 7th. An illustration of the procedure used for each pair of groups can be seen in Figure 1. To control for catching bias (Biro & Post, 2008), we caught all fish in one tank with a single sweep of the net and then split the individuals randomly between two arenas. This was repeated for a second tank to make up two new groups of eight. The 16 new groups were filmed in a random order, two pairs of groups at a time, in four

identical 550mm diameter circular white arenas filled to 3cm water depth. Care was taken to ensure uniform lighting across all arenas. Each group was initially placed in the middle of the arena in a white ring with 12cm diameter for two minutes, which was lifted as filming started, and was then filmed for 10 minutes at 25 frames per second. Videos were tracked in IDTracker (Pérez-Escudero, Vicente-Page, Hinz, Arganda, & de Polavieja, 2014) from the one minute mark until the end, and the median global alignment over this time period was used as each group's score (see below). These ranks were then used to pair the groups in the next round.

The sorting continued for 12 rounds, and was repeated for a total of three times (each instance referred to from here as a 'replicate'), each with an independent collection of 128 guppies. The time between rounds varied between one and 10 days, with a mean of 2.1 days (see Supplementary Information A).

[FIGURE 2 AROUND HERE]

The median of the global alignment over time was used as the sorting measure. This is a standard measure for characterising the directional co-ordination of animals (Kotrschal et al., 2018), also known as polarisation (Couzin et al., 2002) or the order parameter (Vicsek & Zafeiris, 2012). The global alignment is defined as the normalised sum of the eight unit vectors that characterise the directions of motion. This measure is equal to one if all fish have the same orientation, and decreases as they become less aligned. It has recently been shown that familiar groups of guppies are slightly more aligned than non-familiar groups in open field tests, an effect attributed to the attention required to assess unfamiliar conspecifics (Davis, Lukeman, Schaerf, & Ward, 2017). We believe that with our method, this effect should be uniform across all groups, as all guppies would need to assess four unfamiliar or semi-familiar conspecifics each round. In addition, the average time between rounds was much shorter than the approximate 12 days required in guppies for developing familiarity (Griffiths & Magurran, 1997).

As illustrated in Figure 2, the global alignment measure is affected by both alignment and cohesion, which are in general difficult to separate (Perna, Grégoire, & Mann, 2014). To independently

measure the cohesion of the groups, we identified the fish that were exploring the arena together in a 'subgroup' for each frame of the video. Any pair of fish that were within 100mm of each other (approx. four body lengths) was counted as part of the same subgroup. On each frame, we identified the number of fish in the largest subgroup, and took the mean of this number over the trial as a measure of cohesion (group size). The median speed over all data points, using one data point per fish per frame, was used to measure the activity of the group.

Finding repeatability from group consistency

The group consistency for a given global measure is defined here as the Spearman rank correlation of the measure between rounds. For example, the group consistency of global alignment is the rank correlation between the 16 global alignments in the current round and the 16 global alignments in the previous round. This quantifies how similar the rankings of the groups are before and after exchanging half of their members. Because this is based on rankings, it is not influenced by factors affecting the behaviour of all groups over time in the same direction (e.g. reduced activity due to habituation). Note that when adjacently ranked groups are paired to form two new groups, the new groups are assigned the identities of the two old groups randomly, since both are composed in an identical way. Under a null hypothesis in which the 16 group rankings are entirely randomised (i.e. if there is no consistency in group-level behaviour over time) the group consistency lies between −0.50 and 0.50 with a probability of 95% for a single round. Trends in group consistency were tested with a linear mixed-effect model, with the square root of the round number and the number of days since the previous round as fixed effects, and replicate as a random effect. Trends in the measures themselves were tested in the same way, with the rank of the group as an additional fixed effect.

It is reasonable to assume that consistent differences between groups are due to consistent differences in some individual trait, usually quantified by repeatability (Bell, Hankison, & Laskowski, 2009). The group consistency of global alignment was therefore fitted to a theoretical model (Szorkovszky et al., 2017), in order to estimate the repeatability of the underlying individual

trait. The expected trend in the group consistency depends on the within-individual variation relative to the between-individual variation, as well as on how the group's outcome is related to its phenotypic composition . In the model, the repeatability R is tuned by adjusting the ratio of the among-individual variance to the within-individual variance. In the limit that R is zero (i.e. there are no consistent differences between individuals), then the group rankings are expected to change randomly at each round and hence the group consistency varies around zero. For small but non-zero R, the group consistency is expected to be above zero on average. For larger R, the group consistency increases to larger values over successive rounds. In the limit that R approaches one (i.e. differences between individuals account for all variation), the consistency of the rankings rapidly approaches one as sorting progresses. Using maximum-likelihood estimation, the simulated R that best fits the outcome can therefore be chosen (Szorkovszky et al., 2017).

We modelled the global alignment of a group separately as a function of the mean, maximum, or variation of phenotypes in the group. For each model, and for each value of R, we performed 1000 simulations of the experiment. The log-likelihood for a given round was then calculated by comparing the experimental group consistency to a histogram of group consistency from the simulations. This was then summed for all rounds of the experiment for the total likelihood of the combination of model and repeatability (see Supplementary Information B).

Subgroup properties

The global measurements above characterise the average alignment, cohesion and activity of all eight individuals. To get a more detailed picture of how the guppies shoal, we define more measurements limited to subgroups (i.e. ≤ 8 guppies in close proximity). The mode of the main subgroup size was recorded in every two-second period (50 frames) to create a coarse-grained time series for each trial. By analysing this time series, probabilities of this main subgroup increasing or decreasing in size were calculated (see Supplementary Information C). We also calculated properties of the subgroups depending on their size. The mean distance from the centre of the subgroup (subgroup radius) was used as a local aggregation measure. Subgroup alignment was

defined within subgroups similarly to the global alignment, and subgroup speed was defined as the median speed of the centre of the subgroup.

Locomotion and interactions

Ultimately, the group level properties of animal groups emerge from differences in the locomotion of individuals within them. Using the fine-detailed tracking data, we can measure how their locomotion varied according to whether they formed part of a higher ranked aligned group, versus a lower ranked aligned group. Guppies swim with intermittent locomotion, with burst and glides motion typical of many species of fish (Weihs, 1974). Many of these bursts are accompanied by a change in angle prior to the increase in speed (Herbert-Read et al., 2017). The discrete nature of these bursts allows us to measure the timings, magnitudes of speed change, and turning angles during this intermittent locomotion (see Supplementary Information F). We can also ascertain the influence of the social environment on this motion by assessing how these decisions are affected by the location or direction of near neighbours.

Turning angles were compared against four potential directional influences. The heading of a fish's nearest neighbour and the mean heading of the group were used as alignment influences. Similarly, the direction to the nearest neighbour and to the centre of the group were used as attraction influences. For each of these four influences, the correlation between the turning angle and the angle of this influence was used as a measure of its strength within each trial.

Body sizes were also obtained from the videos using IDTracker, and adjusted for small differences in lighting between and within arenas (see Supplementary Information G). The sizes were then tested against group ranking, controlling for differences between replicates, to test whether body size influenced the sorting procedure.

Statistics

Trends in the behaviours that were quantified for each group were tested using a linear mixed-effect model (LMM). Group ranking, round number, and the number of days since the previous round

were used as fixed effects. Replicate was used as a random effect. All measures were checked for normality of residuals. Effect sizes are reported as t-statistics. All analysis was performed in MATLAB R2014b.

Results

[TABLE 1 AROUND HERE]

Group-level measures

After 12 rounds of sorting, the top 8 sorted groups, as well as being more aligned, formed tighter groups and moved faster than the bottom 8 groups. Sample videos of top, middle and bottom-ranking groups after the end of sorting are available (see Supplementary Information H). Table 1 shows that group rank had a significant effect on all of our measurements of collective behaviour (see Supplementary Information I for full details of statistics). Figure 3 (a-c) shows the changes in the per-trial mean global alignment, the mean size of the largest group and the median speed over the course of the sorting procedure, averaged over all three replicates. The clearest pattern is a decrease in all three measures over time. During the early rounds of sorting, the guppies are unfamiliar with the test arena, and group tightly while actively exploring the novel environment, leading to a high global alignment, group size and speed. In subsequent rounds, the guppies dispersed more widely around the arena and moved more slowly in both the top and bottom eight groups. The group rankings according to global alignment were highly positively correlated with the rankings according to mean group size and speed. The mean Spearman rank correlation coefficient between mean global alignment and mean size of the largest group was $\rho = 0.72$ ($P < 0.001$), while the mean Spearman rank correlation coefficient between mean global alignment and median speed was $\rho = 0.52$ ($P = 0.02$).

[FIGURE 3 AROUND HERE]

Finding repeatability from group consistency

Even though the groups were mixed every round, groups showed high positive consistency for their global alignment, group size and speed between rounds (Figure 3 d-f). The group consistency of global alignment increased over subsequent rounds (LMM: $t = 3.1$, $P = 0.006$, $df = 30$), but was not influenced by the time between rounds ($t = 0.1$, $P = 0.9$, $df = 30$). However, when a longer time had passed since the previous round, the group consistency of the average speed decreased ($t = -2.0$, $P = 0.05$, $df = 30$).

Another way of showing that the sorting procedure increased the group consistency of the global alignment is by considering the average change in group ranking. For a null hypothesis of the groups changing randomly between rounds, the average change is 5.25 for a 16-group experiment. In our experiment, the average change in ranking according to global alignment, after fitting to a linear trend, decreased from 4.1 to 3.2 between the second and 12th rounds (Figure 4, $t = -2.4$, $P = 0.02$, $df = 30$). Therefore, as adjacent groups became more similar in their phenotypic composition, mixing adjacent groups had a smaller effect on behaviour over time.

[FIGURE 4 AROUND HERE]

By comparing the observed changes in rank to simulations based on simple models that relate group composition to performance (Szorkovszky et al., 2017), we estimated the repeatability of the individual behaviour governing the global alignment. A conservative estimate was obtained using a model in which individuals' contributions to the global alignment were additive and each varied around its respective phenotype. This gave an estimate of $R = 0.43 \pm 0.03$ (S.E.) for the repeatability (see Supplementary Information B). This is close to the figures previously found for sociability ($R = 0.46$) and activity ($R = 0.38$) in female guppies (Brown & Irving, 2014). If we instead assumed in the simulations that global alignment depends on leader/follower relationships, where one or a small number of individuals lead the rest of the group, the best fit to the experimental data was obtained using a higher value of R. The limiting case, where the global alignment depends on the maximum or minimum individual phenotype, led to an estimate for R above 0.7 (see Supplementary Information B).

Subgroup properties

As sorting progressed, subgroups became more likely to break up and, when containing at least six members, became less likely to increase in size (see Supplementary Information D). Higher ranked groups were more likely to increase in size, and less likely to break apart (LMM: $P < 0.001$ for all group sizes larger than 2).

[FIGURE 5 AROUND HERE]

By separating all of the data into subgroups and comparing subgroups of the same number of fish, we could control for differences in average subgroup size. Larger groups moved slower, were less aligned and occupied larger areas (Fig. 5). Higher ranked groups were faster, more aligned and more compact across all subgroup sizes. We found a negative correlation between subgroup speed and subgroup radius when using the average of each trial (N=8, Spearman $\rho = -0.31$, $P < 0.001$), despite the fact that subgroup radius increased with higher subgroup speed within a trial (see Supplementary Information E).

Locomotion and interactions

The burst patterns of the guppies depended on the position of the nearest neighbour, as shown in Figure 6 (a-c), and on round number, time between rounds and group ranking as shown in Table 1. A guppy's speed over time was mainly regulated by the speed 'minima' at which the bursts were made, and the acceleration during these bursts. On average, guppies retained a higher speed when their nearest neighbours were at moderate distances ($\approx 40 - 100$mm). Individuals in higher ranked groups had higher average burst speeds and higher accelerations. While the mean time between bursts was also dependent on nearest neighbour distance as shown, and increased over the experiment, there was no discernible difference between differently-ranked groups.

[FIGURE 6 AROUND HERE]

Another component of burst-and-glide motion is the directional change before bursts. The average turning angles as a function of the four potential directional influences are shown in Figure 6 (d-g).

In each case, the average turning angle was maximal when the direction of influence (φ) was close to 90 degrees. Hence, for each of these four influences, the correlation between the turning angle and sin(φ) provided a useful measurement of the strength of that influence within each trial. The linear mixed-effect models for these measures are shown in Table 1, alongside the four global-level measures. Alignment-based correlations decreased significantly over subsequent rounds, while attraction-based correlations increased. The top ranking groups showed an increased response to the average position and the average orientation of all fish in the trial, as well as the orientation of the nearest neighbour. The direction to the nearest neighbour, however, had a reduced influence in higher-ranked groups. Notably, for only one measure (attraction to the group centre) is the effect of group ranking in the same direction as the effect of round number.

The mean, maximum, minimum and standard deviation in body sizes were also ranked across groups for each round. None of these ranks were significantly correlated with the global alignment rank. Therefore, body size did not significantly influence the sorting procedure (see Supplementary Information G).

Discussion

The consistency of the group rankings increased over the sorting process, clearly supporting our prediction of repeatable variation in shoaling behaviour across groups. The mixing of similar groups did not influence behaviour enough to substantially alter the group rankings, compared to mixing of dissimilar groups. This implies that the alignment of a guppy shoal is driven by a repeatable behavioural phenotype rather than temporal factors such as familiarity.

The differences between groups with high and low alignment can be explained by a combination of differences in activity and attraction to conspecifics, or sociability, both of which are well-established repeatable traits in guppies (Budaev, 1997; Burns, 2008; Brown & Irving, 2014). This is reflected in measurements at all three of the levels we examined. Individuals in top ranked groups were more active, as shown by a higher average speed as well as higher speed at bursting times.

Individuals in top-ranked groups also formed larger and more stable subgroups and were more likely to stay close to the centre of these, as shown by the smaller spatial spread of subgroups, increased turning responses to the group centre and lower likelihood of groups breaking up. Interestingly, individuals in low-ranking groups, rather than showing a reduced response to all social stimulus, instead showed a relative increase in responses to their nearest neighbours' positions. This could indicate that sociability differences may be not in the strength, but in the selectivity of interactions (Strandburg-Peshkin et al., 2013; Jiang et al., 2017). Although top-ranked groups show increased alignment with the nearest neighbour, this may be a side-effect of higher activity (i.e. both individuals closely following the group trajectory) rather than from a direct social response to the neighbour. Our analysis of shoaling groups provides a more natural quantification of shoaling tendency than standard sociability assays, in which individuals are assessed in how they approach a shoal, generally situated behind a barrier (Budaev, 1997; Brown & Irving, 2014). Notably, such individual assays evaluate the tendency to join a group, whereas from studying the changes in subgroup sizes we can conclude that there is even greater variation in the tendency to leave groups. With our method, individuals can be assigned scores based on the rank of their final group, although this limits the number of possible levels to the number of groups.

For most measurements we analysed, the effect of group ranking (i.e. increased group alignment) is opposite to the effect of round number (Table 1). A potential explanation for this observation is that higher ranked groups habituate slower to the repeated assays. Individual differences in habituation have been found in threespine sticklebacks (*Gasterosteu aculeatus*), separate from other personality differences (Bell & Peeke, 2012). Notably, in our experiment the turning response to the centre of the group, which has been found to be a relatively accurate predictor of anti-predator responses in guppies (Kimbell & Morrell, 2015), increased with both time and group ranking. This indicates that potential habituation differences cannot explain the observed social differences between low-ranked and high-ranked groups. To fully investigate the habituation effect, additional control replicates are needed in which groups are either kept unsorted or shuffled randomly for the same number of

rounds.

We also quantified the repeatability *R* of behaviour by fitting a simulated experiment to the data. Although *R* is generally quantified by multiple assays on individuals (Réale et al., 2007; Bell, Hankison & Laskowski, 2009), we were able to estimate *R* using the changes in group rankings (Szorkovszky et al., 2017). Our conservative repeatability estimate was similar to what has been reported for repeatability of activity and sociability as measured in assays on individual female guppies (Brown & Irving, 2014). This result is consistent with the conclusion that sociability and activity are the main contributors to variation in shoaling behaviour as tested here. Additionally we find a positive correlation between activity and sociability. This may contrast with what has recently been shown in threespine sticklebacks, where more social fish were found to be less active in both individual and collective settings (Jolles, Boogert, Sridhar, Couzin, & Manica, 2017). The variation in sociability is consistent with that found between guppies from high and low-predation environments respectively (Herbert-Read et al., 2017). This predation study used a similar open field test in groups of eight, but found no effect of predation on overall activity. This suggests that activity and sociability do not form a syndrome in guppies, and our choice of alignment as the sorting measure introduced the positive correlation.

Guppies in the wild self-assort by social tendency, although whether this is due to active decisions to associate with similar conspecifics or due to a passive mechanism is still an open question (Croft et al., 2005). Our results suggest that this self-assortment can emerge passively from the simple differences we have identified. In an open environment, the individuals with high sociability could maintain links with each other (as observed in networks) by staying as close as possible to the centre of the group, while individuals with low sociability form smaller and more numerous cliques, by staying close to their nearest neighbours. Such predictions can be verified by using simulated models of collective motion (Eriksson, Jacobi, Nyström, & Tunstrøm, 2010; Sumpter, 2010).

The method that we have devised is a novel way of analysing the variation in behaviours, such as shoaling, that are difficult to quantify with individual assays. Results on the underlying quantitative

394 structure of the trait rely on simulations, which may require strong assumptions in some cases.
395 However, work can be done to extend the validity of our approach to more general situations. For
396 instance, in the simulations we have assumed various functions for how the group-level property
397 emerges from individual behaviours (Farine, Montiglio & Spiegel 2015), and then fitted a single
398 parameter to estimate the repeatability. For cases in which the group-level measurement is (either
399 explicitly or equivalently) a sum of individual behaviours, it is also possible to adopt an indirect
400 genetic effects formalism (Moore, Brodie & Wolf 1997) in the simulated model. For example,
401 interactions between phenotypes may be modelled by a single interaction coefficient (Bijma 2014)
402 which is expected to affect the sorting dynamics. Hence, this parameter may also be estimated in
403 addition to the repeatability. As interactions between multiple conspecifics are difficult to decouple
404 using regression methods (Bijma 2014), our approach may provide a fruitful alternative.

405 We have demonstrated strong and repeatable variation in shoaling behaviour that persists during
406 repeated mixing of groups, and which may explain common patterns of self-assortment. This
407 laboratory setup provides the opportunity for a number of follow-up experiments uncovering
408 collective behaviour. For instance, directed breeding of sorted individuals will allow the heritability
409 of the sorted traits to be determined (Drent, van Oers, & van Noordwijk, 2003), and if heritable,
410 assays of subsequent generations will reveal which traits coevolve with these aspects of social
411 behaviour. Such further tests will provide experimental data on the ecological function and
412 evolution of variation in shoaling behaviour.

Acknowledgements

The authors would like to thank Damien Farine for assistance as editor, as well as multiple anonymous referees for helpful comments on previous versions of this article. This work was supported by the Knut and Alice Wallenberg Foundations, grant number 102 2013.0072.

References

Aplin, L. M., Farine, D. R., Morand- Ferron, J., Cole, E. F., Cockburn, A., & Sheldon, B. C. (2013). Individual personalities predict social behaviour in wild networks of great tits (*Parus major*). *Ecology Letters*, 16(11), 1365-1372.

Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, 77(4), 771–783.

Bell, A. M., & Peeke, H. V. (2012). Individual variation in habituation: behaviour over time toward different stimuli in threespine sticklebacks (gasterosteus aculeatus). *Behaviour*, 149(13-14), 1339–1365.

Bijma, P. (2014). The quantitative genetics of indirect genetic effects: a selective review of modelling issues. *Heredity*, 112(1), 61.

Biro, P. A., & Post, J. R. (2008). Rapid depletion of genotypes with fast growth and bold personality traits from harvested fish populations. *Proceedings of the National Academy of Sciences*, 105(8), 2919–2922.

Brown, C., & Irving, E. (2014). Individual personality traits influence group exploration in a feral guppy population. *Behavioral Ecology*, 25(1), 95–101.

Budaev, S. (1997). “Personality” in the guppy (*Poecilia reticulata*): A correlational study of

438 exploratory behavior and social tendency. *Journal of Comparative Psychology*, 111, 399–411.

439 Burns, J. G. (2008). The validity of three tests of temperament in guppies (*Poecilia reticulata*).
440 *Journal of Comparative Psychology*, 122(4), 344.

441 Couzin, I. D., Krause, J., James, R., Ruxton, G. D., & Franks, N. R. (2002). Collective memory and
442 spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1), 1–11.

443 Croft, D., James, R., Ward, A., Botham, M., Mawdsley, D., & Krause, J. (2005). Assortative
444 interactions and social networks in fish. *Oecologia*, 143(2), 211–219.

445 Davis, S., Lukeman, R., Schaerf, T. M., & Ward, A. J. W. (2017). Familiarity affects collective
446 motion in shoals of guppies (*Poecilia reticulata*). *Royal Society Open Science*, 4(9). doi:
447 10.1098/rsos.170312

448 Dingemanse, N. J., Kazem, A. J., Réale, D., & Wright, J. (2010). Behavioural reaction norms:
449 animal personality meets individual plasticity. *Trends in Ecology & Evolution*, 25(2), 81–89.

450 Dingemanse, N. J., & Araya-Ajoy, Y. G. (2015). Interacting personalities: Behavioural ecology
451 meets quantitative genetics. *Trends in Ecology & Evolution*, 30(2), 88–97.

452 Drent, P. J., van Oers, K., & van Noordwijk, A. J. (2003). Realized heritability of personalities in
453 the great tit (*Parus major*). *Proceedings of the Royal Society of London B: Biological Sciences*,
454 270(1510), 45–51.

455 Dugatkin, L. A., & Godin, J.-G. J. (1992). Predator inspection, shoaling and foraging under
456 predation hazard in the Trinidadian guppy, *Poecilia reticulata*. *Environmental Biology of Fishes*,
457 34(3), 265–276.

458 Dyer, J. R., Croft, D. P., Morrell, L. J., & Krause, J. (2009). Shoal composition determines foraging
459 success in the guppy. *Behavioral Ecology*, 20(1), 165–171.

460 Eriksson, A., Jacobi, M. N., Nyström, J., & Tunstrøm, K. (2010). Determining interaction rules in
461 animal swarms. *Behavioral Ecology*, 21(5), 1106–1111.

462 Farine, D. R., & Whitehead, H. (2015). Constructing, conducting and interpreting animal social
463 network analysis. *Journal of Animal Ecology*, 84(5), 1144-1163.

464 Farine, D. R., Montiglio, P.-O., & Spiegel, O. (2015). From individuals to groups and back: The
465 evolutionary implications of group phenotypic composition. *Trends in Ecology & Evolution*,
466 30(10), 609–621.

467 Farr, J. A. (1975). The role of predation in the evolution of social behavior of natural populations of
468 the guppy, *Poecilia reticulata (pisces: Poeciliidae)*. *Evolution*, 29(1), 151–158.

469 Foster, W., & Treherne, J. (1981). Evidence for the dilution effect in the selfish herd from fish
470 predation on a marine insect. *Nature*, 293, 466–467.

471 Griffiths, S. W., & Magurran, A. E. (1997). Familiarity in schooling fish: how long does it take to
472 acquire? *Animal Behaviour*, 53(5), 945–949.

473 Hamilton, W. D. (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, 31(2), 295–
474 311.

475 Herbert-Read, J. E., Krause, S., Morrell, L. J., Schaerf, T. M., Krause, J., & Ward, A. J. W. (2013).
476 The role of individuality in collective group movement. *Proceedings of the Royal Society of*
477 *London. Biological Sciences*, 280(1752), 20122564.

478 Herbert-Read, J. E., Rosén, E., Szorkovszky, A., Ioannou, C. C., Rogell, B., Perna, A., . . . Krause,
479 J. (2017). How predation shapes the social interaction rules of shoaling fish. *Proceedings of the*
480 *Royal Society B: Biological Sciences*, 284(1861).

481 Ioannou, C. C., Guttal, V., & Couzin, I. D. (2012). Predatory fish select for coordinated collective
482 motion in virtual prey. *Science*, 1218919.

483 Jiang, L., Giuggioli, L., Perna, A., Escobedo, R., Lecheval, V., Sire, C., ... Theraulaz, G. (2017).
484 Identifying influential neighbors in animal flocking. *PLoS Computational Biology*, 13(11),
485 e1005822.

Jolles, J. W., Boogert, N. J., Sridhar, V. H., Couzin, I. D., & Manica, A. (2017). Consistent individual differences drive collective behaviour and group functioning of schooling fish. *Current Biology*, 27(18), 2862-2868.

Killen, S. S., Marras, S., Nadler, L., & Domenici, P. (2017). The role of physiological traits in assortment among and within fish shoals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1727), 20160233.

Kimbell, H. S., & Morrell, L. J. (2015). 'Selfish herds' of guppies follow complex movement rules, but not when information is limited. *Proceedings of the Royal Society B: Biological Sciences*, 282(1816).

King, A. J., Williams, L. J., & Mettke-Hofmann, C. (2015). The effects of social conformity on Gouldian finch personality. *Animal Behaviour*, 99, 25-31.

Kotrschal, A., Szorkovszky, A., Romenskyy, M., Perna, A., Buechel, S. D., Zeng, H.-L., . . . Kolm, N. (2018). Brain size does not impact shoaling dynamics in unfamiliar groups of guppies (*Poecilia reticulata*). *Behavioural processes*, 147, 13–20.

Krause, J., Butlin, R. K., Peuhkuri, N., & Pritchard, V. L. (2000). The social organization of fish shoals: a test of the predictive power of laboratory experiments for the field. *Biological Reviews*, 75(4), 477–501.

Krause, J., & Ruxton, G. D. (2002). *Living in groups*. Oxford University Press.

Landeau, L., & Terborgh, J. (1986). Oddity and the "confusion effect" in predation. *Animal Behaviour*, 34(5), 1372–1380.

Magnhagen, C., & Staffan, F. (2005). Is boldness affected by group composition in young-of-the-year perch (*Perca fluviatilis*)? *Behavioral Ecology and Sociobiology*, 57(3), 295–303.

Magurran, A. E., & Seghers, B. H. (1994). Predator inspection behaviour covaries with schooling tendency amongst wild guppy, *Poecilia reticulata*, populations in Trinidad. *Behaviour*, 128(1),

121–134.

Metcalfe, N. B., & Thomson, B. C. (1995). Fish recognize and prefer to shoal with poor competitors. *Proceedings of the Royal Society of London B: Biological Sciences*, 259(1355), 207–210.

Michelena, P., Jeanson, R., Deneubourg, J. L., & Sibbald, A. M. (2010). Personality and collective decision-making in foraging herbivores. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1684), 1093-1099.

Miller, N., Garnier, S., Hartnett, A. T., & Couzin, I. D. (2013). Both information and social cohesion determine collective decisions in animal groups. *Proceedings of the National Academy of Sciences*, 110(13), 5263–5268.

Moore, A.J., Brodie, E.D. & Wolf, JB (1997). Interacting phenotypes and the evolutionary process: I. direct and indirect genetic effects of social interactions. *Evolution* 51: 1352–1362

Pélabon, C., Larsen, L.-K., Bolstad, G., Viken, Å., Fleming, I., & Rosenqvist, G. (2014). The effects of sexual selection on life-history traits: an experimental study on guppies. *Journal of Evolutionary Biology*, 27(2), 404–416.

Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., & de Polavieja, G. G. (2014). idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods*, 11(7), 743–748.

Perna, A., Grégoire, G., & Mann, R. P. (2014). On the duality between interaction responses and mutual positions in flocking and schooling. *Movement Ecology*, 2(1), 22.

Pike, T. W., & Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biology Letters*, 6(4), 466.

Pike, T. W., Samanta, M., Lindström, J., & Royle, N. J. (2008). Behavioural phenotype affects social interactions in an animal network. *Proceedings of the Royal Society of London B: Biological*

Sciences, 275(1650), 2515–2520.

Pruitt, J. N., & Goodnight, C. J. (2014). Site-specific group selection drives locally adapted group compositions. *Nature*, 514(7522), 359.

Réale, D., Reader, S. M., Sol, D., McDougall, P. T., & Dingemanse, N. J. (2007). Integrating animal temperament within ecology and evolution. *Biological Reviews*, 82(2), 291–318.

Rosenthal, S. B., Twomey, C. R., Hartnett, A. T., Wu, H. S., & Couzin, I. D. (2015). Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proceedings of the National Academy of Sciences*, 112(15), 4690–4695.

Seghers, B. H. (1974). Schooling behavior in the guppy (*Poecilia reticulata*): an evolutionary response to predation. *Evolution*, 28(3), 486–489.

Shalizi, C. R., & Thomas, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, 40(2), 211– 239.

Sih, A., Cote, J., Evans, M., Fogarty, S., & Pruitt, J. (2012). Ecological implications of behavioural syndromes. *Ecology Letters*, 15(3), 278–289.

Sih, A., & Watters, J. V. (2005). The mix matters: behavioural types and group dynamics in water striders. *Behaviour*, 142(9-10), 1417–1431.

Silk, J. B. (2007). The adaptive value of sociality in mammalian groups. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 539–559.

Strandburg-Peshkin, A., Twomey, C. R., Bode, N. W., Kao, A. B., Katz, Y., Ioannou, C. C., . . . Levin, S. A. (2013). Visual sensory networks and effective information transfer in animal groups. *Current Biology*, 23(17), R709–R711.

Sumpter, D. J. (2010). Collective animal behavior. Princeton University Press.

Sumpter, D. J., & Pratt, S. C. (2009). Quorum responses and consensus decision making. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1518), 743–

753.

Sundaresan, S. R., Fischhoff, I. R., Dushoff, J., & Rubenstein, D. I. (2007). Network metrics reveal differences in social organization between two fission–fusion species, grevy's zebra and onager. *Oecologia*, 151(1), 140–149.

Szorkovszky, A., Kotrschal, A., Herbert Read, J., Sumpter, D., Kolm, N., & Pelckmans, K. (2017). An efficient method for sorting and quantifying individual social traits based on group-level behaviour. *Methods in Ecology and Evolution*. doi: 10.1111/2041210X.12813

van der Bijl, W., Thyselius, M., Kotrschal, A., & Kolm, N. (2015). Brain size affects the behavioural response to predators in female guppies (*Poecilia reticulata*). *Proceedings of the Royal Society B: Biological Sciences*, 282(1812).

Vicsek, T., & Zafeiris, A. (2012). Collective motion. *Physics Reports*, 517(3), 71–140.

Webster, M. M., & Ward, A. J. (2011). Personality and social context. *Biological Reviews*, 86(4), 759–773.

Weihs, D. (1974). Energetic advantages of burst swimming of fish. *Journal of Theoretical Biology*, 48(1), 215–229.

Weinmann, A., Storath, M., & Demaret, L. (2015). The l^1-potts functional for robust jump-sparse reconstruction. *SIAM Journal on Numerical Analysis*, 53(1), 644–673.

Wolf, M., & Weissing, F. J. (2012). Animal personalities: consequences for ecology and evolution. *Trends in Ecology & Evolution*, 27(8), 452–461.

Appendix A: Time between rounds

[Figure A1 HERE]

Appendix B: Model fitting

To fit the empirical rank transitions to simulations, we used maximum-likelihood estimation. In the sorting simulations, the between-individual variation was kept constant and the within-individual variation σ was varied. This was done for three prototypical models of group behaviour, where the group feature (i.e. the global alignment) depends on the mean, maximum or standard deviation of the individual phenotypes comprising the group (Szorkovszky et al., 2017). For a given experimental round t, the group repeatability $C(t)$ was compared against likelihood histograms from 1000 iterations of each model and σ parameter. This was done for every round after the second, treating each as independent. For a single replicate, the likelihood for a value of σ over the twelve rounds was then calculated as

$$L(\sigma)=\sum_{t=2}^{12}\log l_\sigma(t,C(t))$$

where $l_\sigma(t,c)$ is the likelihood of the correlation in round t being c, based on histograms obtained from performing several simulations with within-individual variation parameter σ. The replicates were then combined by adding the log-likelihood for all three.

[FIGURE A2 AROUND HERE]

The between-individual variation was set to one, so the individual repeatability R was then defined as (Bell, Hankison, & Laskowski, 2009)

$$R=\frac{1}{1+\sigma^2}$$

This resulted in a curve for each of the three models, showing the relative likelihoods of the

parameter R, as shown in Figure A2. The lowest estimate was for the 'mean' model with $R \approx 0.43$. A standard error of 0.03 was determined using the asymptotic normal approximation, by fitting a fourth-order polynomial to the combined log-likelihood function and calculating the second derivative from the fitted coefficients.

Appendix C: Subgroups

For each point in time, subgroups were defined by counting any pair of fish < 100mm apart as part of the same subgroup. We denote the position vector of a fish with label i as $\vec{r}_i(t)$ and define its instantaneous velocity as

$$\vec{v}_i(t) = \vec{r}_i(t) - \vec{r}_i(t-1)$$

where t > 1 denotes the frame of the video. The centroid of the subgroup G with n members is defined simply as

$$\vec{r}_G(t) = \sum_{i \in G} \vec{r}_i(t)$$

If the membership of subgroup G was kept constant from the previous frame, the instantaneous speed of subgroup G was then defined as the speed of its centroid

$$S_G(t) = \left| \vec{r}_G(t) - \vec{r}_G(t-1) \right|$$

The radius of subgroup G is defined as the average distance from the centre

$$R_G(t) = \sum_{i \in G} \left| \vec{r}_i(t) - \vec{r}_G(t) \right|$$

The subgroup alignment was given by the length of the sum of the normalised velocities of the members, divided by the number of members n_G

$$A_G(t) = \frac{1}{n_G} \left| \sum_{i \in G} \frac{\vec{v}_i(t)}{\left| \vec{v}_i(t) \right|} \right|$$

Appendix D: Subgroup size transitions

The mode of the main subgroup size was recorded in every two second period (50 frames) to create a coarse-grained time series for each trial. These time series were collated according to group ranking and sorting round. Figure A3 shows the transition probabilities calculated from these time series, for groups in the top 8 and bottom 8 of the ranking. The more aligned groups are less likely to decrease in main subgroup size, and more likely to increase. The effect sizes for group size decreases are consistently larger (Table A1).

[FIGURE A3 AROUND HERE]

Appendix E: Subgroup speed and radius

Within each trial, the Spearman rank correlation was calculated between sub- group speed and radius (mean distance from group centroid) for all frames where all fish were in the same subgroup. This correlation averaged $\rho = 0.18$ with a standard deviation of 0.18 ($\rho > 0 : t = 26.2$, $P < 0.001$). The correlation between the per-trial medians of these measures took the opposite sign ($\rho = -0.31$, $P < 0.001$). This can also be seen the two-dimensional frequency distributions, separated according to rank and time, as shown in Figure A4.

[FIGURE A4 AROUND HERE]

Appendix F: Burst and glide analysis

Speeds were calculated from the first derivatives of the x and y time series, then smoothed using a third-order Savitzky-Golay filter. A peak-finding algorithm was then used to find local maxima and minima, with each set constrained to be at least one third of a second apart. The maxima and minima were then paired to determine accelerations during each burst (speed at maximum minus

speed at prior minimum), as well as subsequent glide times (time of minimum minus time of prior maximum, limited to four seconds to control for inactive periods).

The heading angles of the fish were also calculated from the first derivatives of the x and y time series, then unwrapped. An L1-Potts functional was used to detect jumps upward or downward in this angle (Weinmann, Storath, & Demaret, 2015), which were recorded as turning decisions.

Appendix G: Body sizes

Eight body sizes were estimated for each trial, accounting for changes in apparent size between the middle and edge of the arena. Sizes were then corrected using a linear mixed-effect model. The fixed effect of time accounted for growth of the fish over the experiment, and a fixed effect of group ranking was used to test for a body size effect on sorting. Random effects were replicate and arena, accounting for different ages and lighting conditions, respectively. The resulting corrected body sizes are shown in Figure A5. After correction, the body sizes are approximately normally distributed, with no discernible difference in the distribution between the top and bottom four groups. From the model, the fitted size difference between the top and bottom group is 4.7% of the residual standard deviation (group rank effect $P = 0.20$). From this, we can conclude that the sorting did not operate on the body size.

[FIGURE A5 AROUND HERE]

Appendix H: Videos

Three one-minute videos of sorted groups in the open field arena are available in the Online Supplementary Information. These are taken from the 7th minute of trials, taken 7 days after the final sorting round in replicate 2. Corresponding smoothed plots of global alignment vs time are shown in Figure A6. The shaded regions correspond to the location of the videos.

[FIGURE A6 AROUND HERE]

Appendix I: Full statistics

For the measures presented in the main text, full statistics including intercepts, estimates and standard errors (SE) are contained in Table A2. Refer to main text for full description of each measure. The number in parentheses denotes the degrees of freedom of the error (DFE). Residuals are plotted in Figure A7.

[FIGURE A7 AROUND HERE]

Appendix J: Single-group plots

An alternative version of Figure 3(a-c) from the main text, where measures are plotted against group rank instead of round, can be seen in Figure A8. To see the consistency between the last two rounds of sorting at the group level, we show changes in global alignment after exchanging members as a reaction norm plot (Dingemanse, Kazem, Réale, & Wright, 2010) in Figure A9.

[FIGURE A8 AROUND HERE]

[FIGURE A9 AROUND HERE]

Tables and Figures

Table 1: Summary of fixed effect sizes for all measured variables

Measure	Round		Days		Rank	
	t	*P*	*t*	*P*	*t*	*P*
Global align	-27.1	<0.001	5.4	<0.001	39.5	<0.001
Speed	-6.9	<0.001	3.7	<0.001	13.1	<0.001
Group size	-11.3	<0.001	2.6	0.010	21.9	<0.001
Radius (N=8)	3.4	0.001	-1.0	0.331	-7.7	<0.001
Speed minimum	-6.8	<0.001	4.9	<0.001	12.6	<0.001
Acceleration	-0.6	0.518	-2.2	0.025	7.5	<0.001
Glide time	9.1	<0.001	-3.5	0.001	0.1	0.910
Group align	-14.8	<0.001	1.0	0.310	21.5	<0.001
Group attract	3.3	0.001	0.5	0.637	11.7	<0.001
N.n. Align	-13.3	<0.001	1.0	0.339	19.3	<0.001
N.n. Attract	12.2	<0.001	-2.7	0.008	-6.8	<0.001

Fixed effect sizes for round number, days between rounds and group rank are shown as t-statistics with associated P-values. For all tests above the error degree of freedom $df = 524$. Group-level measures are Global align: median global alignment; Speed: median speed; Group size: mean size of largest subgroup; Radius: median average distance to global centroid when all eight are in the same subgroup. Individual-level measures are Speed minimum: mean speed at beginning of burst; Acceleration: mean minimum to maximum speed change during burst; Glide time: mean time between maximum and next minimum. The final four measures are correlations of individuals' turning angles with the sine of angular influences. Group align: mean heading of the group; Group attract: direction of the centre of the group; N.N. align: heading of the nearest neighbour; N.N. attract: direction to the nearest neighbour. See Supplementary Information I for full details.

Table A1: Effect sizes for subgroup transitions

Measure	*df*	Round *t*	Round *P*	Day *t*	Day *P*	Rank *t*	Rank *P*	Ratio
Pup (N=2)	412	-0.70	0.485	-2.54	0.011	2.89	0.004	-0.24
Pup (N=3)	517	-1.17	0.242	0.75	0.454	7.45	<0.001	-0.16
Pup (N=4)	523	0.33	0.744	0.20	0.842	6.01	<0.001	0.05
Pup (N=5)	523	0.04	0.967	0.01	0.995	7.15	<0.001	0.01
Pup (N=6)	523	-2.34	0.019	0.98	0.328	4.96	<0.001	-0.47
Pup (N=7)	523	-5.34	<0.001	0.99	0.324	7.32	<0.001	-0.73
Pdn (N=3)	517	2.43	0.016	-0.39	0.698	-8.02	<0.001	-0.30
Pdn (N=4)	523	4.96	<0.001	-1.60	0.110	-9.66	<0.001	-0.51
Pdn (N=5)	523	5.09	<0.001	-0.42	0.675	-11.25	<0.001	-0.45
Pdn (N=6)	523	5.47	<0.001	-2.38	0.018	-9.25	<0.001	-0.59
Pdn (N=7)	523	8.64	<0.001	-2.64	0.009	-14.08	<0.001	-0.61
Pdn (N=8)	523	9.61	<0.001	-3.09	0.002	-12.05	<0.001	-0.80

Fixed effect sizes for round number, days between rounds and group rank, shown as t-statistics, and associated P-values, for the probabilities of transitions up (Pup) and down (Pdn) in the main subgroup size. The right-most column shows the round number effect size (*t*) divided by the group rank effect size. If this is greater than zero, both effects are in the same direction.

728

729

730

731

732 Table A2: Full statistics of measures presented in the main text.

733

	Estimate	SE	*t*	*P*
Global align				
Intercept	0.872	0.022	40.305	<0.001
Round	-0.076	0.003	-27.075	<0.001
Rank	-0.015	0.000	-39.513	<0.001
Day	0.005	0.001	5.398	<0.001
Speed				
Intercept	59.209	3.614	16.385	<0.001
Round	-2.864	0.415	-6.894	<0.001
Rank	-0.726	0.055	-13.124	<0.001
Day	0.503	0.135	3.711	0.000
Group size				
Intercept	7.407	0.125	59.390	<0.001
Round	-0.313	0.028	-11.264	<0.001
Rank	-0.081	0.004	-21.900	<0.001
Day	0.023	0.009	2.602	0.010
Radius (*N*=8)				
Intercept	54.425	2.598	20.948	<0.001
Round	1.963	0.578	3.396	0.001
Rank	0.592	0.077	7.659	<0.001
Day	-0.182	0.187	-0.973	0.331
Group align				
Intercept	0.410	0.020	20.638	<0.001
Round	-0.044	0.003	-14.790	<0.001
Rank	-0.008	0.000	-21.467	<0.001
Day	0.001	0.001	1.017	0.310
Group angle				
Intercept	0.398	0.029	13.706	<0.001
Round	0.013	0.004	3.323	0.001
Rank	-0.006	0.001	-11.723	<0.001
Day	0.001	0.001	0.472	0.637

N.n. Align				
Intercept	0.398	0.024	16.471	<0.001
Round	-0.045	0.003	-13.267	<0.001
Rank	-0.009	0.000	-19.257	<0.001
Day	0.001	0.001	0.957	0.339
N.n. Angle				
Intercept	0.179	0.023	7.651	<0.001
Round	0.037	0.003	12.235	<0.001
Rank	0.003	0.000	6.757	<0.001
Day	-0.003	0.001	-2.675	0.008
Speed min				
Intercept	32.608	2.226	14.646	<0.001
Round	-1.838	0.272	-6.767	<0.001
Rank	-0.456	0.036	-12.614	<0.001
Day	0.433	0.089	4.881	<0.001
Acceleration				
Intercept	68.164	2.803	24.316	<0.001
Round	-0.189	0.292	-0.647	0.518
Rank	-0.290	0.039	-7.451	<0.001
Day	-0.214	0.095	-2.247	0.025
Glide time				
Intercept	11.616	0.365	31.813	<0.001
Round	0.496	0.055	9.108	<0.001
Rank	-0.001	0.007	-0.114	0.910
Day	-0.062	0.018	-3.461	0.001

Figure 1: Mixing the first pair of groups in a sorting round. The groups are initially paired according to the previous round's rankings. 1. A random adjacent pair of groups is chosen. 2. Each group is separated into groups of four, and the groups are mixed. 3. The new groups are filmed in two arenas. 4. The new groups are put into tanks, re-numbered in order of filming. Steps 1-4 are repeated until all 16 groups have been mixed and filmed. The videos are then tracked and ranked according to the global alignment.

Figure 2: Examples of low and high global alignment. At a given point in time, the eight fish have positions and orientations denoted by arrows. They form, in arena (a) a large subgroup with low local alignment, in (b) two small subgroups with high local alignment, and in (c) a large subgroup with high local alignment. Only example (c) shows high global alignment.

Figure 3: Changes in group-level measures and group rank consistency during sorting. In panels (a)-(c) the measurements are separated into the top 8 and bottom 8 groups, ranked by the global alignment. In each case, the median (lines) and interquartile range (error bars) of the 24 observations (8 groups, 3 replicates) are shown for each round. Panels (d)-(f) show the group rank consistency of the given measure for each replicate. The dotted line at zero represents the average group consistency for the null hypothesis of random group rankings. Alternative representations of the data in panels (a-c) are available in Supplementary Information J.

Figure 4: Changes in rank over time. Shown is the mean absolute change in rank per round of experiment (red circles) and a linear fit (red line). For comparison is the expected mean absolute change in rank if groups change rank randomly (dotted black line).

Figure 5: Properties of subgroups of different sizes during final round of sorting. (a) Frequency of a fish being a member of a subgroup of a given size for (red) top eight and (blue) bottom eight groups, averaged over time and replicate. The other panels display the (b) speed of the group centre, (c) alignment, and (d) average distance from group centre, for different group sizes. The markers represent the median values, and error bars represent the interquartile range.

Figure 6: Characterisation of observed burst-and-glide motion of individuals as a function of the social environment. Solid lines indicate (a) the mean speed at beginning of burst, defined as a local minimum; and (b) the mean glide time, defined as the time between a local maximum in speed and the next minimum, as a function of the distance to the nearest neighbour. (c) Frequency of nearest neighbour distance over all burst events. (d-e) Mean turning angle as a function of the direction to the nearest neighbour and its relative orientation, when the nearest neighbour is 25-100mm away. (f-g) Mean turning angle as a function of the direction to the average position of all eight individuals, and their average relative orientation, when the average position is < 100mm away. Data is from the final six rounds of sorting, averaged over all replicates. Shaded areas represent ±1 standard error.

Figure A1: Histogram of the number of days between rounds in the experiment.

Figure A2: Log-likelihood estimates of the individual repeatability for each model.

Figure A3: Main subgroup size and transitions. (a) Probability distribution of the main subgroup size N for the top 8 and bottom 8 ranked groups in the first six rounds. (b-c) Box plots of the probability per trial of transitions (b) up and (c) down in group size, for all trials in the first six rounds of sorting. Probabilities are calculated from the average likelihood of the main subgroup size either increasing or decreasing in the following two second period, given a current main subgroup size N. Panels (d-f) are as above but for the final six rounds of sorting.

Figure A4: Two-dimensional histograms of group speed versus group radius when all eight fish were in the same subgroup. Lighter colour indicates higher likelihood. The vertical and horizontal position of the markers denote the mean values for the respective measures.

Figure A5: Histogram of all estimated body sizes in the top and bottom four of the group rankings, in the final four rounds of sorting, after correcting for time, arena and replicate.

Figure A6: Global alignment vs time for the highest, 8th highest and lowest ranked groups (replicate 2). Each point is the median of the surrounding two- second period (51 frames). The shaded regions correspond to the videos in the Online Supplementary Information.

Figure A7: Residuals (vertical axis) plotted against fitted values (horizontal axis) for the linear mixed-effect models.

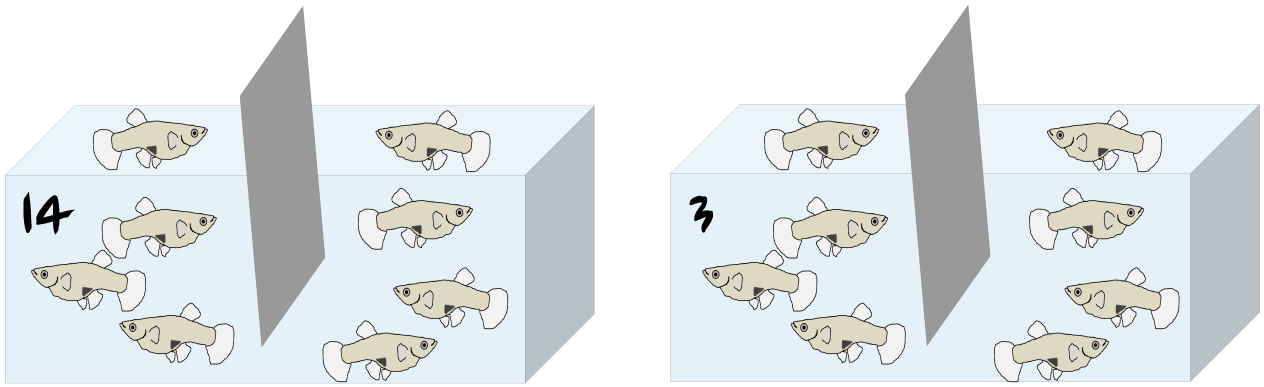
Figure A8: Group-level measures plotted against group ranking for the first round (grey) and the final round (black) for each replicate.

Figure A9: Behavioural reaction norms for the final two rounds of assays. The vertical axis is the global alignment. Each line connects the group measure before and after exchanging four group members. Darker shade indicates higher ranking (by global alignment) before exchanging members.

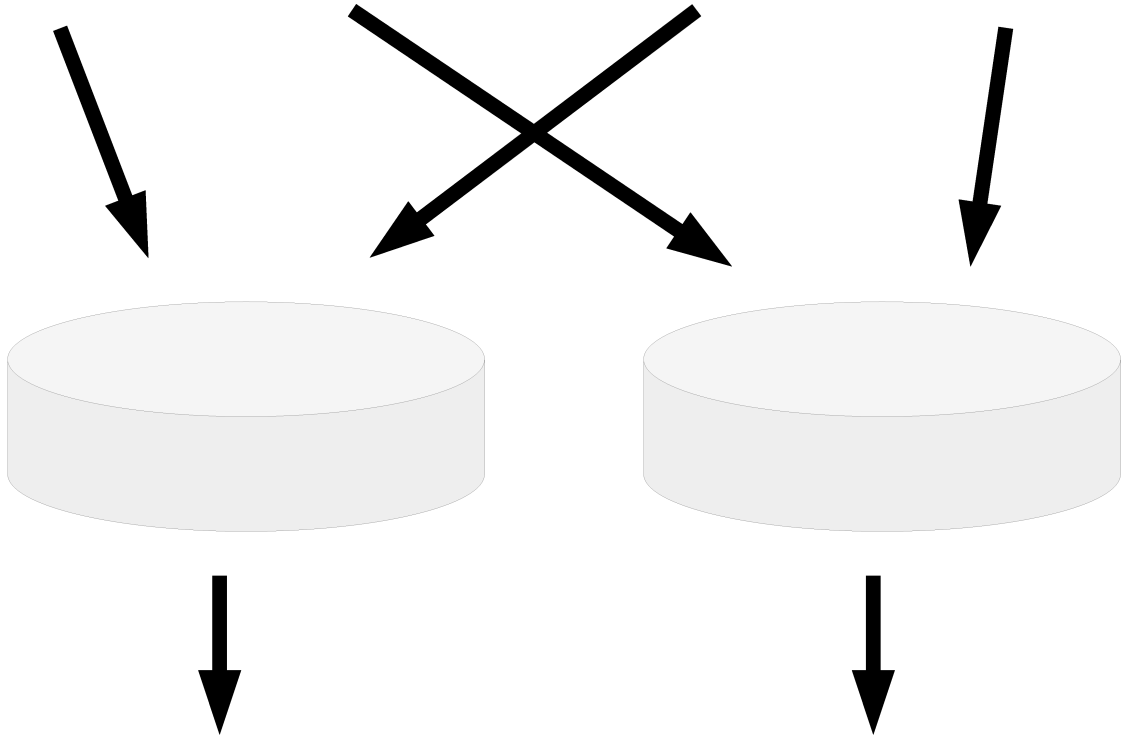
1.

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Tank	5	7	11	2	1	16	14	3	9	4	8	15	12	6	13	10

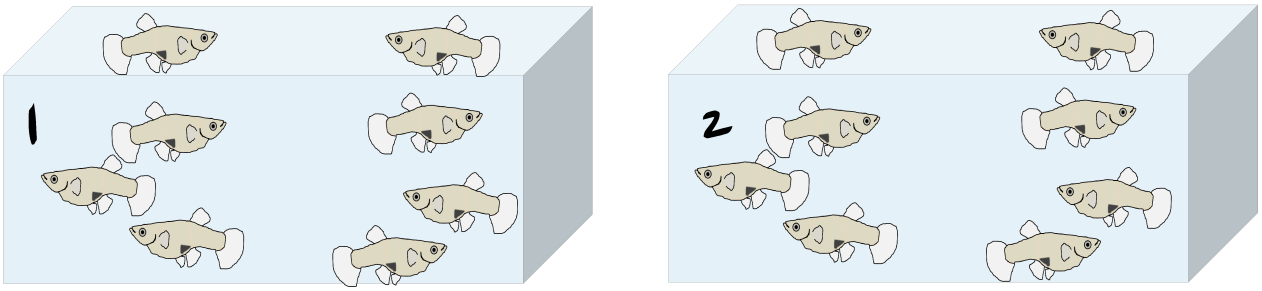
2.

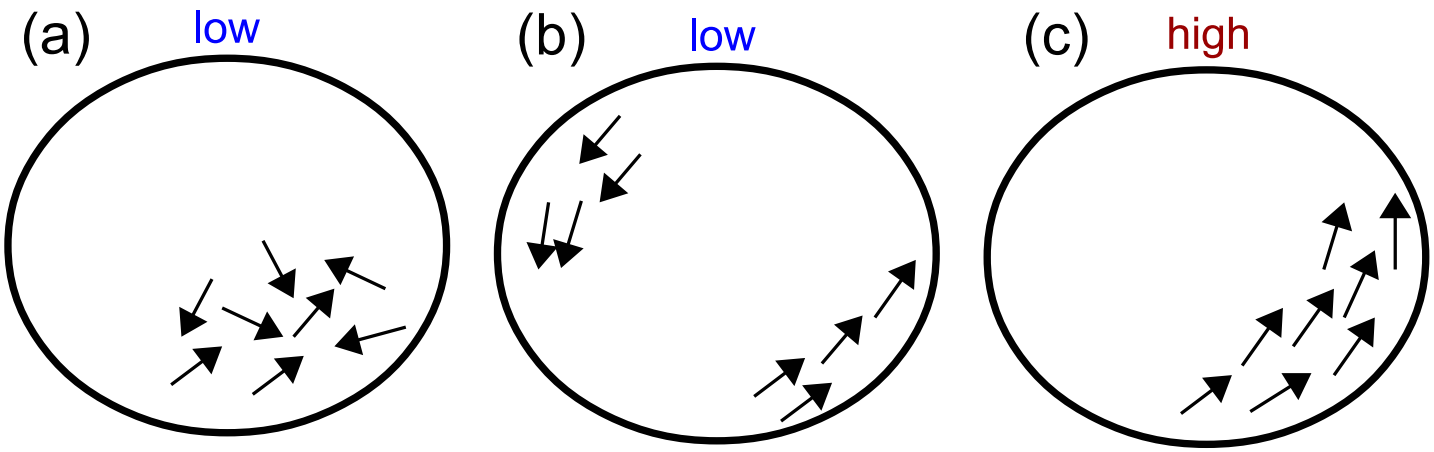


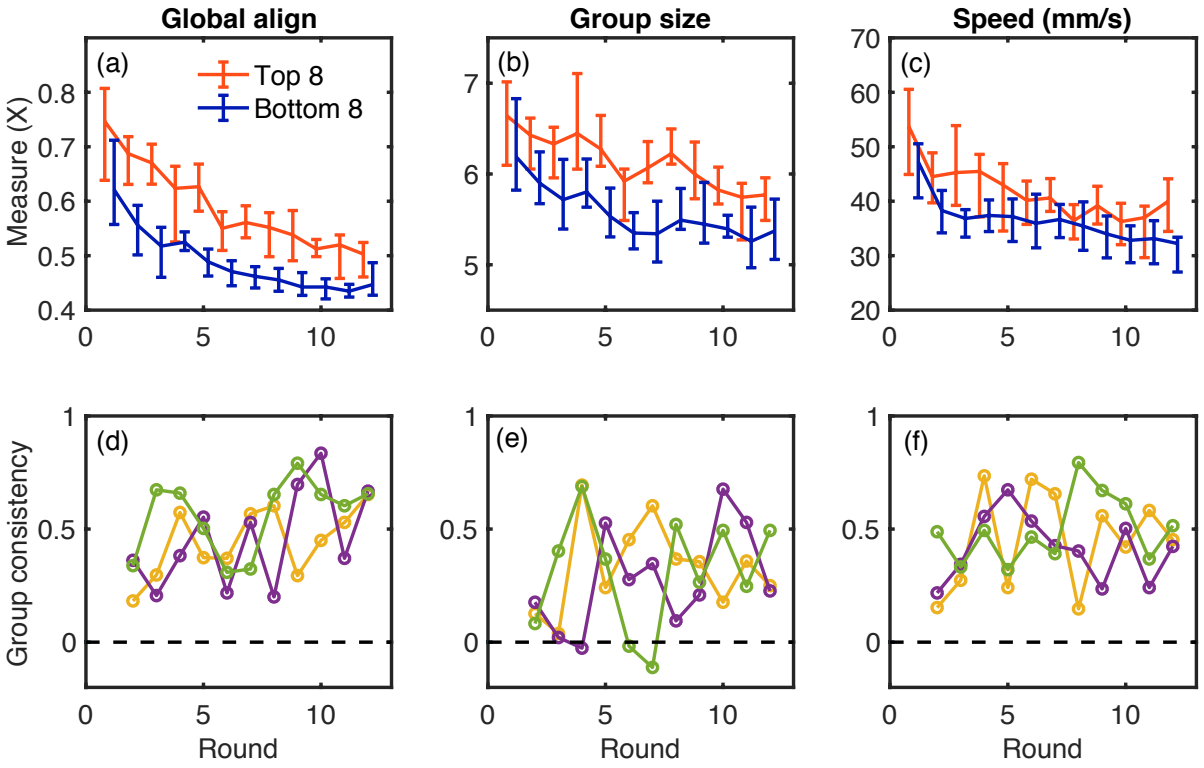
3.

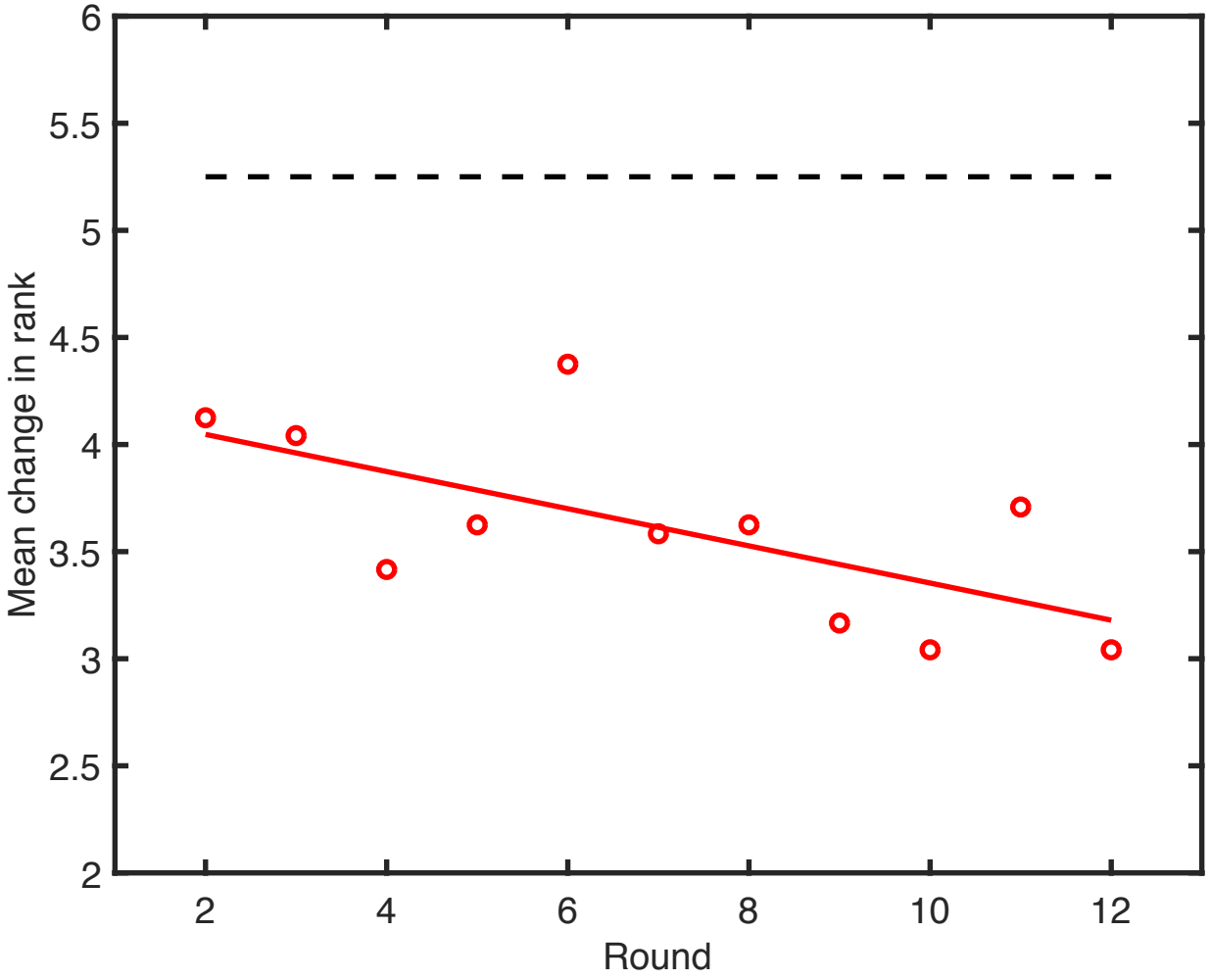


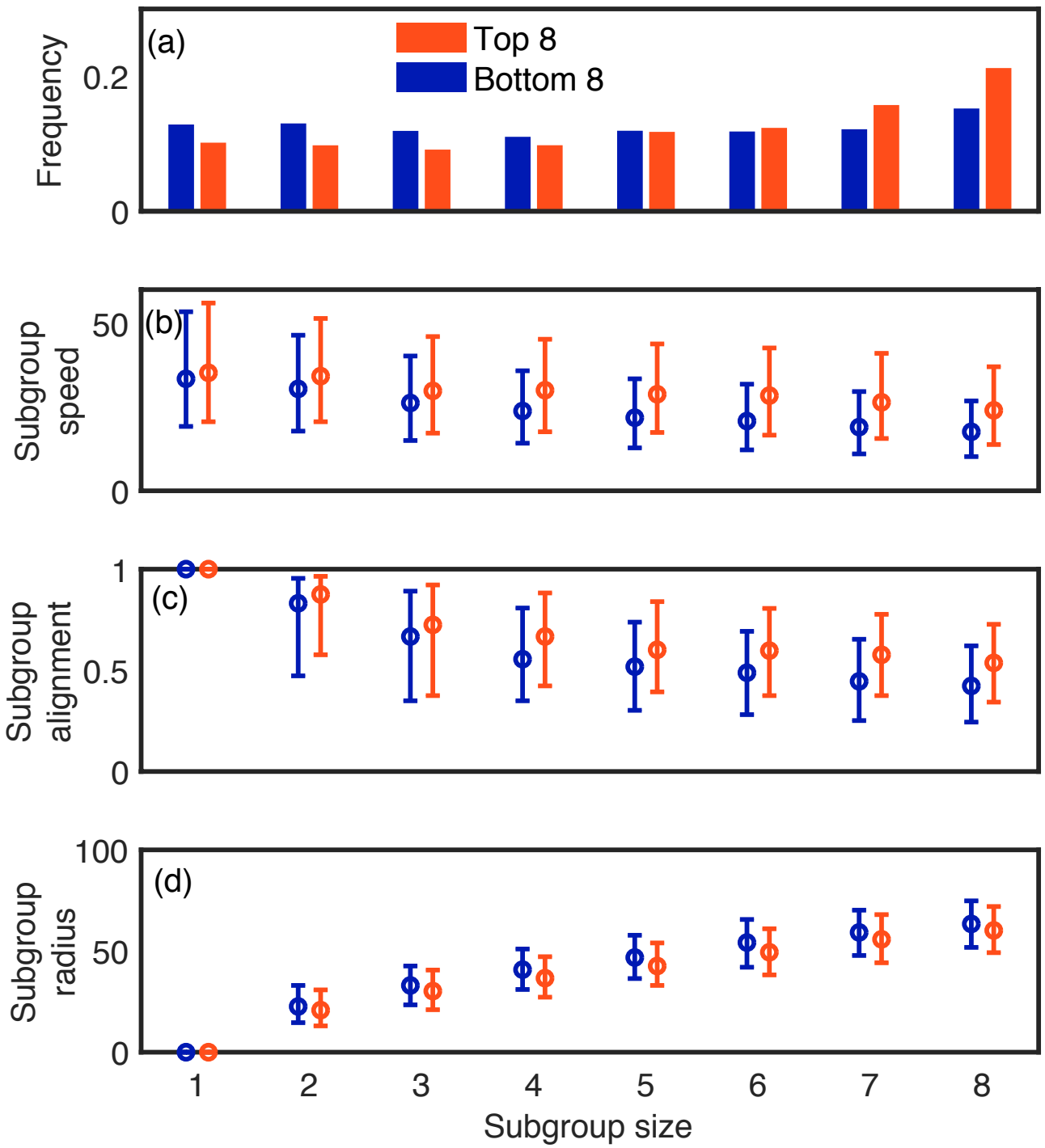
4.

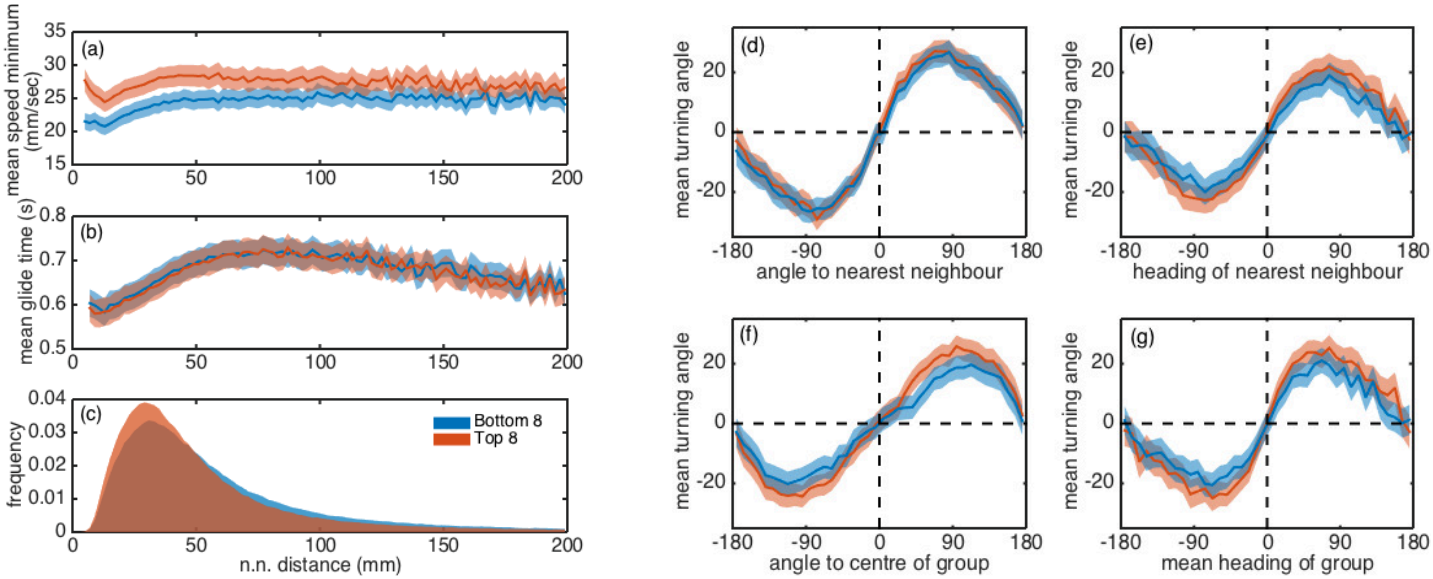


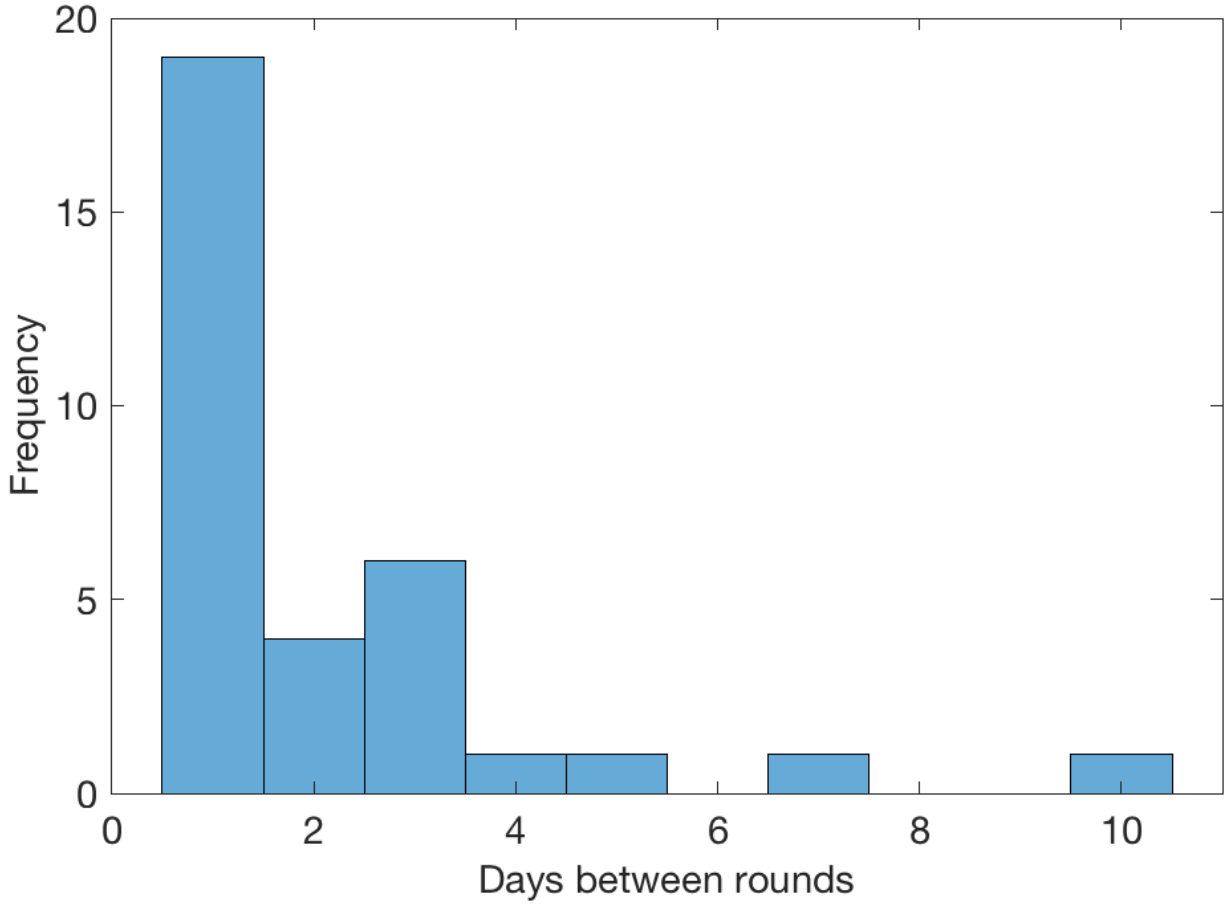

(a)
low
(b)
low
(c)
high

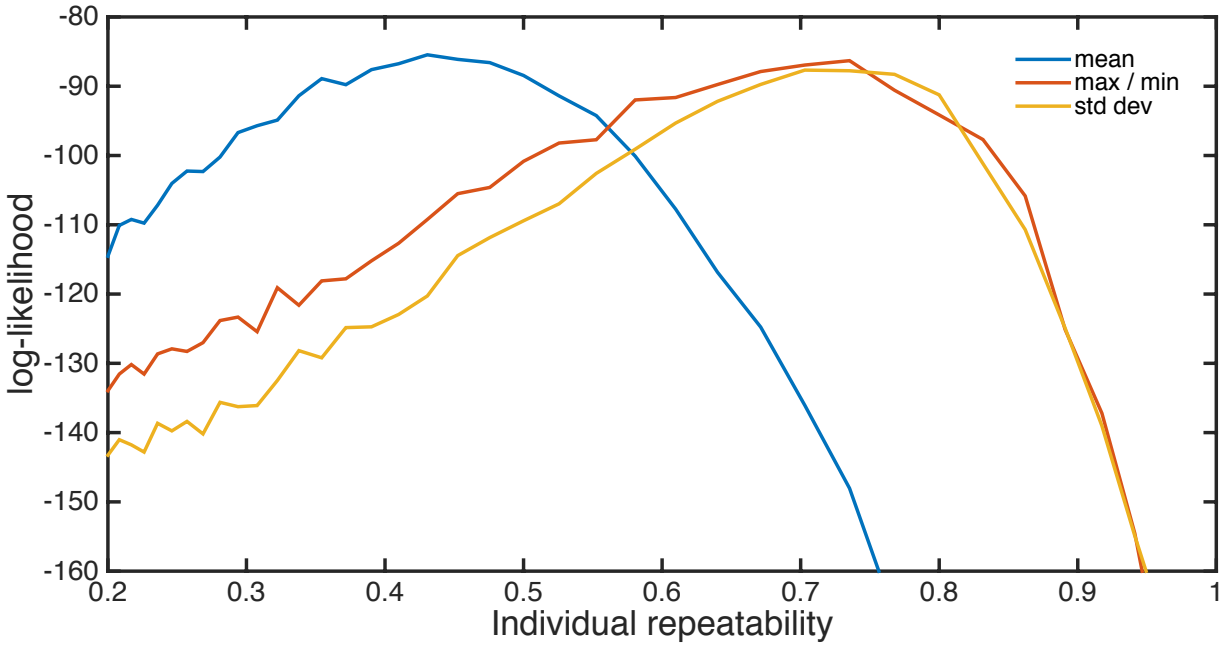

Global align
Group size
Speed (mm/s)
(a)
(b)
(c)
(d)
(e)
(f)
Top 8
Bottom 8
Measure (X)
Group consistency
Round
Round
Round

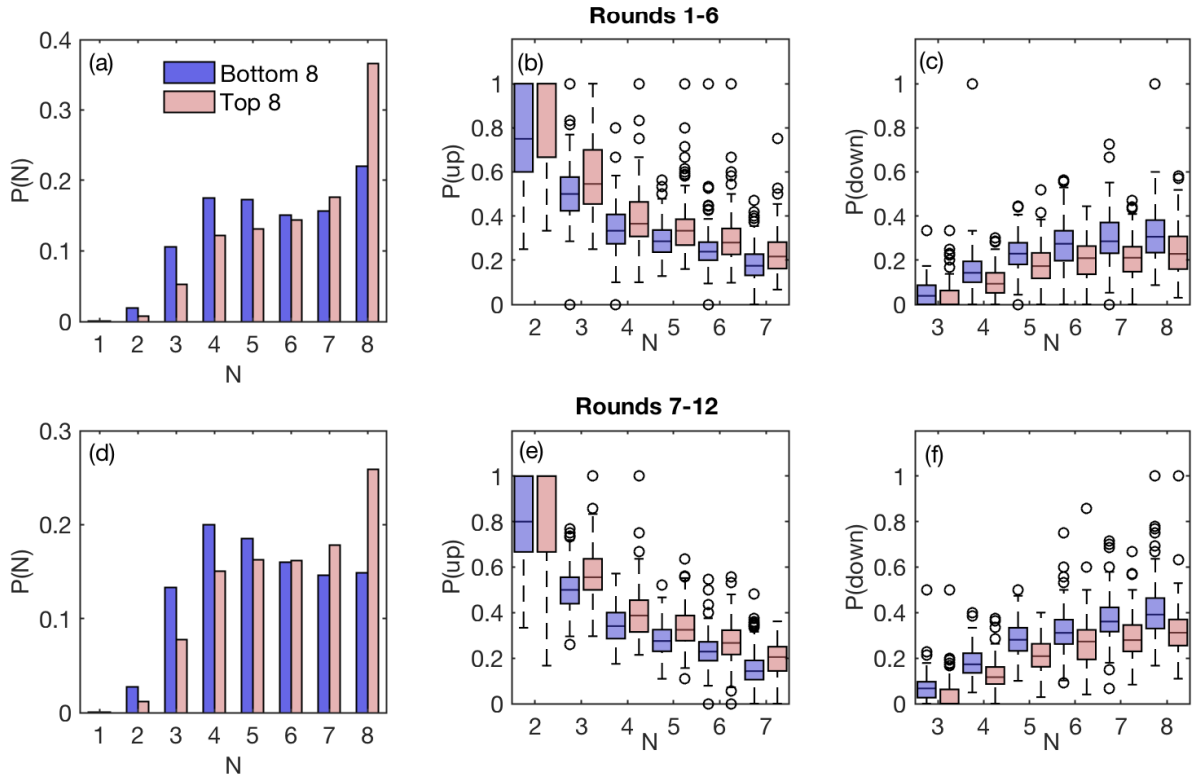

6
5.5
5
4.5
4
3.5
3
2.5
2
Mean change in rank
2
4
6
8
10
12
Round

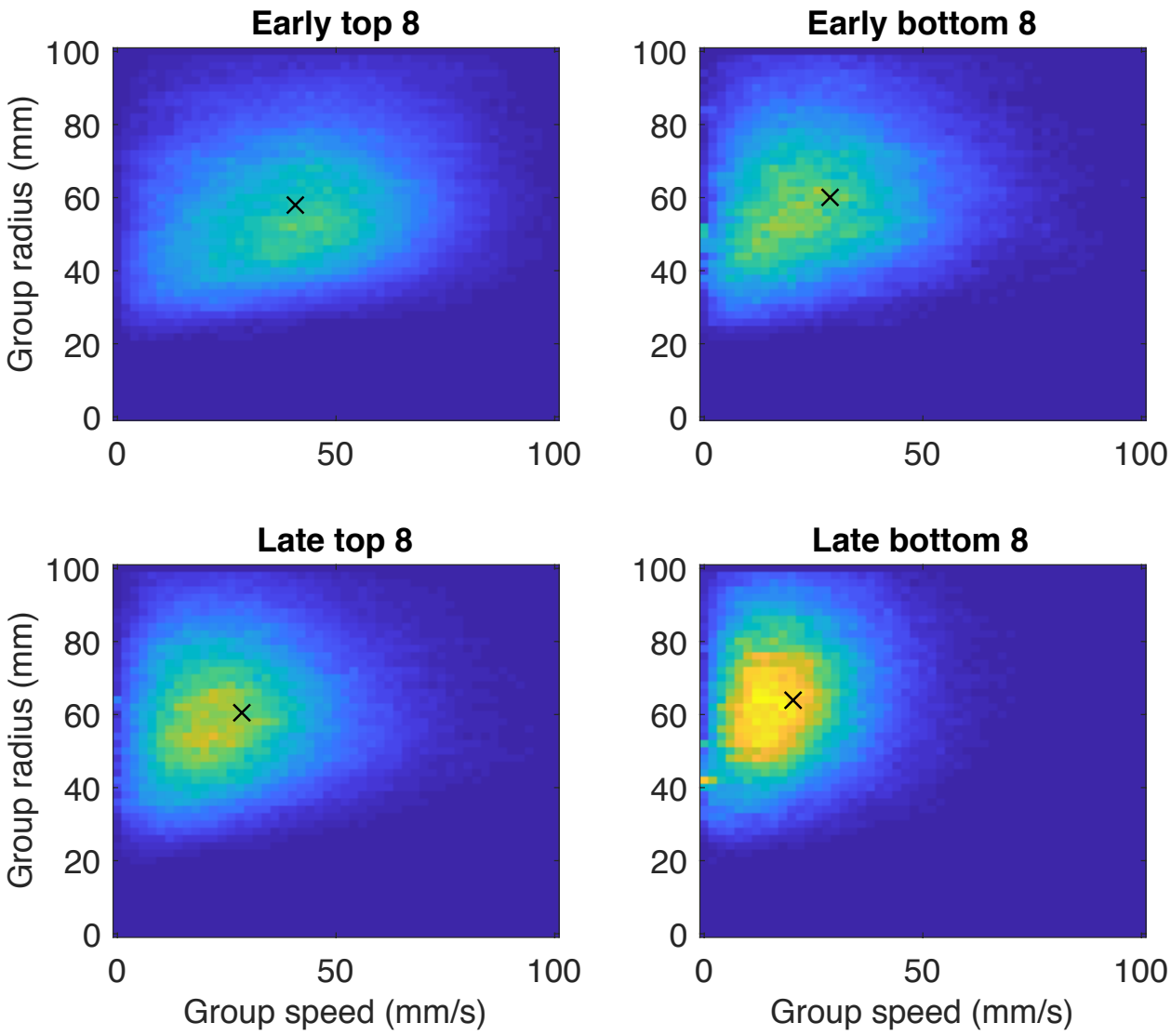

(a)
Top 8
Bottom 8
Frequency
0.2
0
(b)
Subgroup speed
50
0
(c)
Subgroup alignment
1
0.5
0
(d)
Subgroup radius
100
50
0
1
2
3
4
5
6
7
8
Subgroup size

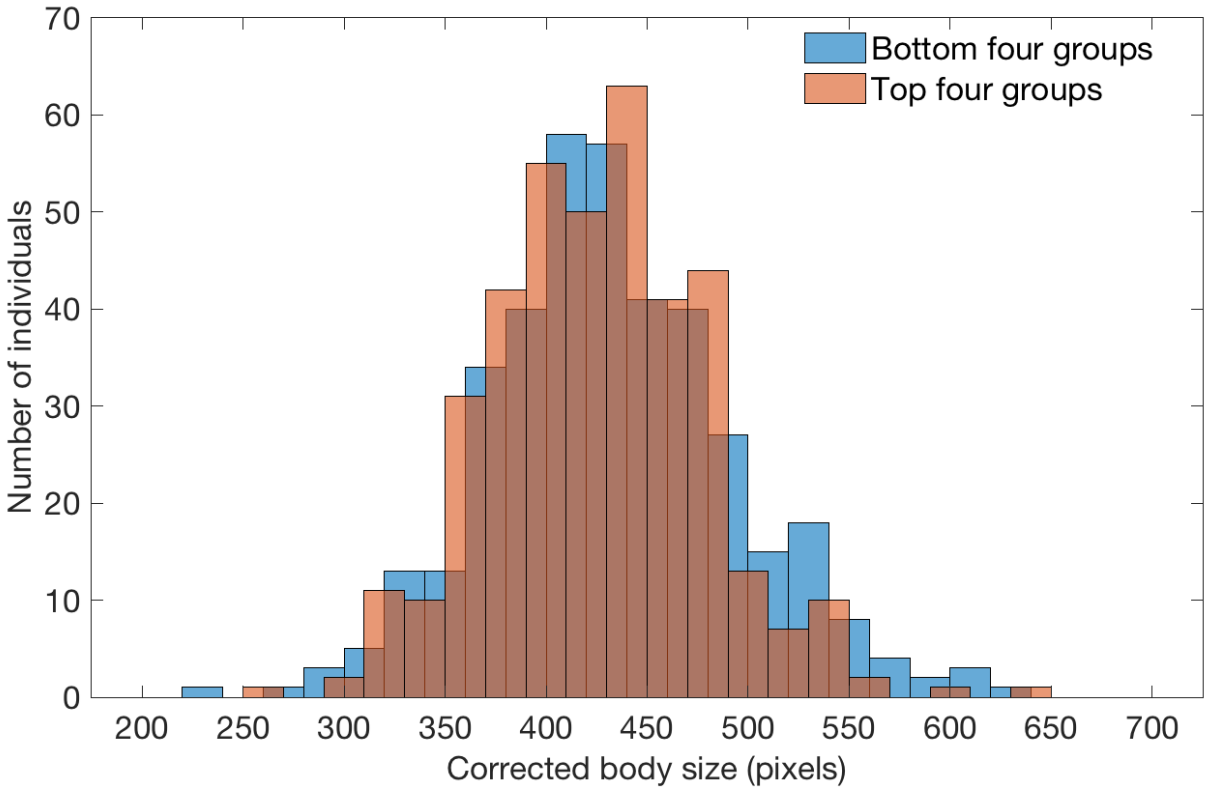

(a)
mean speed minimum (mm/sec)
35
30
25
20
15
0
50
100
150
200
(b)
mean glide time (s)
0.8
0.7
0.6
0.5
(c)
frequency
0.04
0.03
0.02
0.01
0
Bottom 8
Top 8
n.n. distance (mm)
(d)
mean turning angle
20
0
-20
-180
-90
0
90
180
angle to nearest neighbour
(e)
heading of nearest neighbour
(f)
angle to centre of group
(g)
mean heading of group

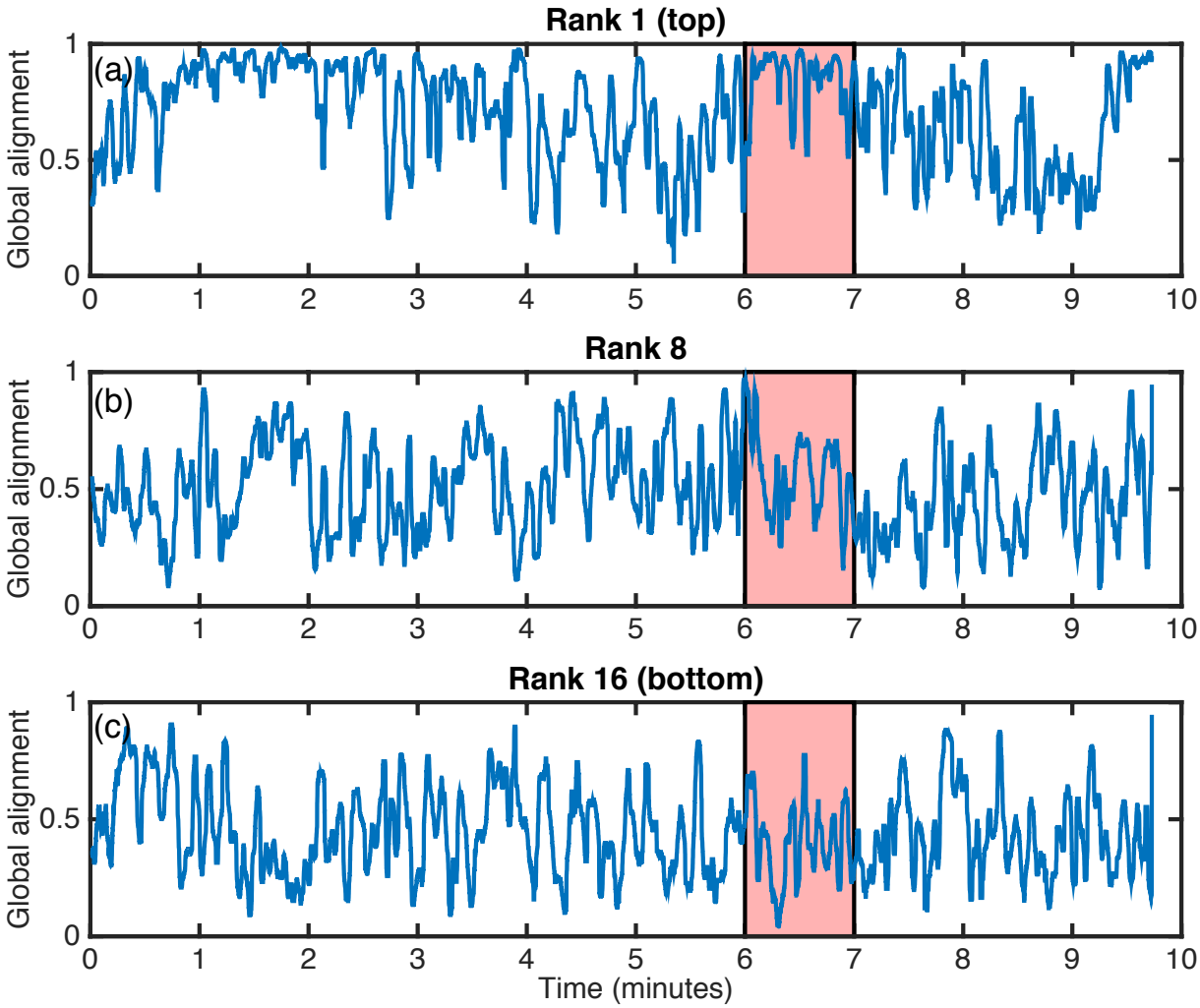

20
15
10
5
0
Frequency
0
2
4
6
8
10
Days between rounds

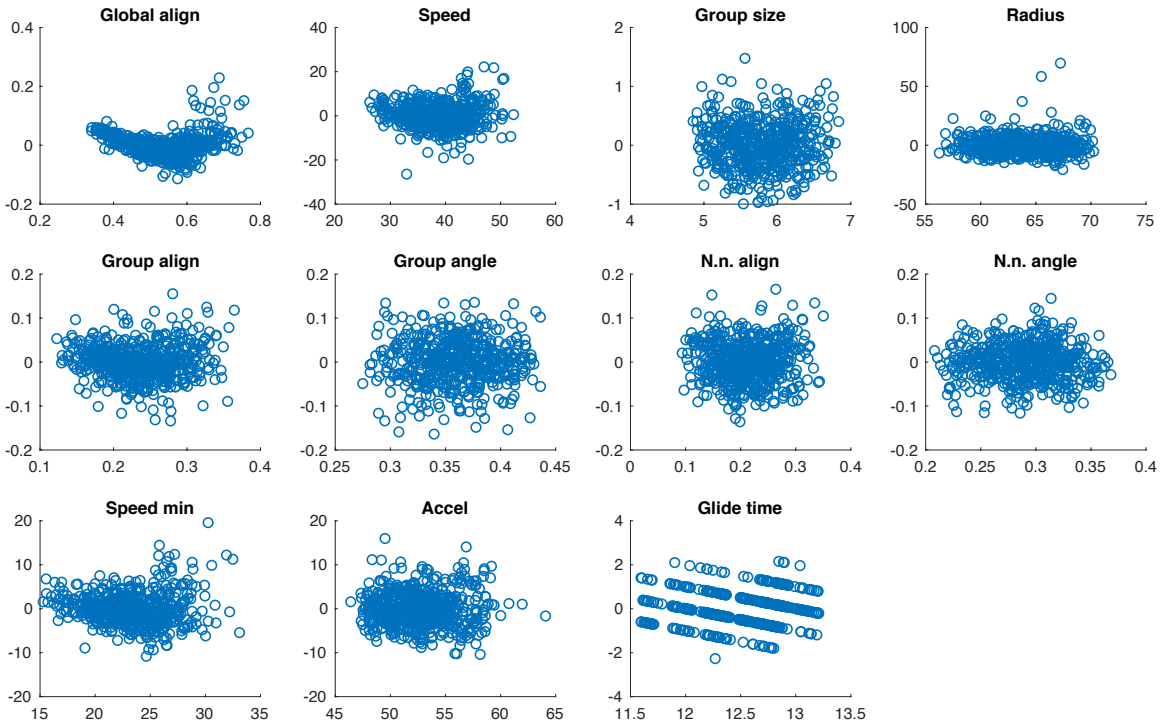

-80
-90
-100
-110
-120
-130
-140
-150
-160
log-likelihood
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Individual repeatability
mean
max / min
std dev

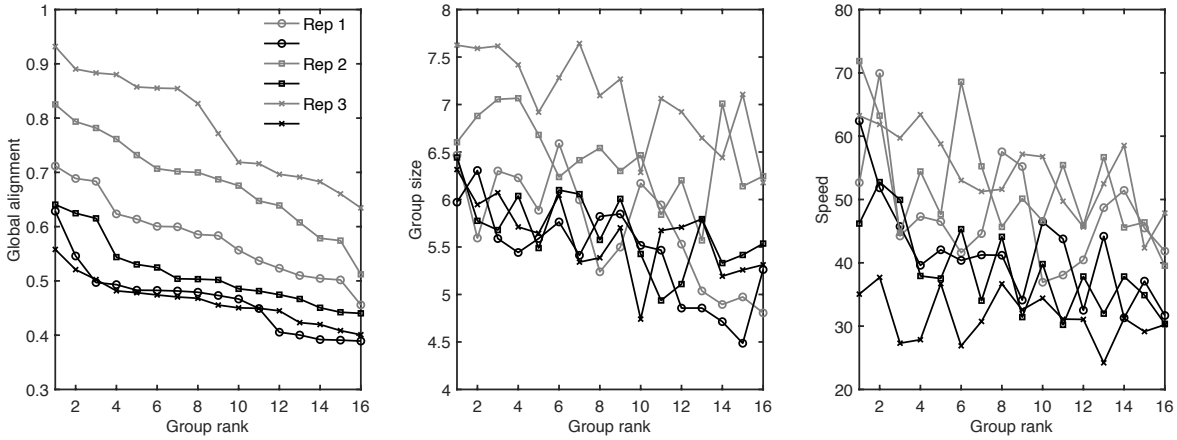

Rounds 1-6
Rounds 7-12
(a)
(b)
(c)
(d)
(e)
(f)
Bottom 8
Top 8
P(N)
P(up)
P(down)
N

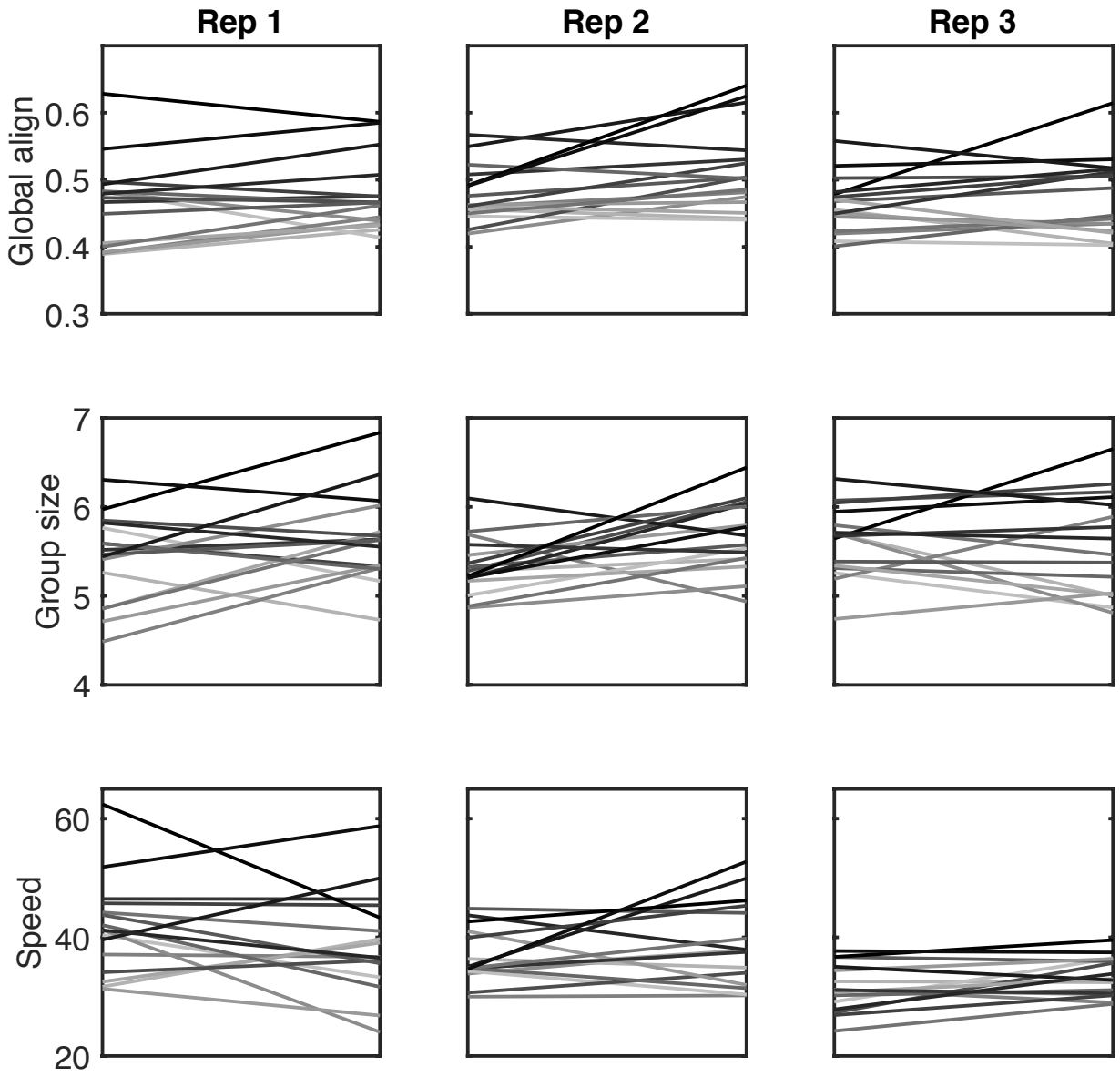

Early top 8
Early bottom 8
Late top 8
Late bottom 8
Group radius (mm)
Group speed (mm/s)
0
20
40
60
80
100
0
50
100


Bottom four groups
Top four groups
Number of individuals
Corrected body size (pixels)
0
10
20
30
40
50
60
70
200
250
300
350
400
450
500
550
600
650
700


Rank 1 (top)
(a)
Global alignment
1
0.5
0
0
1
2
3
4
5
6
7
8
9
10
Rank 8
(b)
Global alignment
1
0.5
0
0
1
2
3
4
5
6
7
8
9
10
Rank 16 (bottom)
(c)
Global alignment
1
0.5
0
0
1
2
3
4
5
6
7
8
9
10
Time (minutes)


Global align
Speed
Group size
Radius
Group align
Group angle
N.n. align
N.n. angle
Speed min
Accel
Glide time


Rep 1
Rep 2
Rep 3
Global alignment
Group size
Speed
Group rank
Group rank
Group rank


Rep 1
Rep 2
Rep 3
Global align
0.6
0.5
0.4
0.3
Group size
7
6
5
4
Speed
60
40
20